UNIVERSIDAD CARLOS III DE MADRID

TESIS DOCTORAL

Essays in the Theory of Repeated Games

Autor:
António Miguel Osório-Costa

Director:
Juan Pablo Rincón-Zapatero

DEPARTAMENTO DE ECONOMÍA

Getafe, Junio de 2010

TESIS DOCTORAL

Essays in the Theory of Repeated Games

Autor: António Miguel Osório-Costa

Director: Juan Pablo Rincón-Zapatero

Firma del Tribunal Calificador:

Firma

Presidente:

Vocal:

Vocal:

Vocal:

Secretario:

Calificación:

Getafe, de Junio de 2010

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I am deeply grateful to my advisor Juan Pablo Rincón-Zapatero for his guidance, encouragement and generous help. This thesis has significantly benefited from his insightful comments and discussions.

The thesis has greatly benefited from useful comments and suggestions from internal and external faculty members, and graduate students. In particular, I want to thank Venkataraman Bhaskar, Antonio Cabrales, Marco Celentani, Luis Corchón, Jacques Crémer, Martin Cripps, Ramiro De Elejalde, M^a Ángeles De Frutos, Klaus Desmet, Mathias Dewatripont, Mikhail Drugov, Jose Luis Ferreira, Drew Fudenberg, Daniel García-González, Penélope Hernández, Ángel Hernando-Veciana, Kenneth Judd, Patrick Legros, David Levine, Albert Ma, Matilde P. Machado, Francisco Marhuenda, Francesc Marmol, Carmelo Nuñez, Wojciech Olszewski, Carlos Ponce, David Rahman, Juan Pablo Rincón-Zapatero, Tomas Sjöström, Andrzej Skrzypacz, Carlos Velasco, Michael Wiper, and Galina Zudenkova.

The research leading to this thesis received the financial support from Fundação para Ciência e Tecnologia grant SFRH/BD/17157/2004 and from the Spanish Ministerio of Ciencia and Innovación Project ECO2008-02358.

Finally, I am most grateful to my family for their encouragement and support. Nothing would be possible without them. I also want to dedicate this thesis to the memory of my father, Evaristo Ferreira da Costa.

SUMMARY

This thesis comprises three essays in economic theory. The first two are in the theory of repeated games. The third is also a theoretical contribution, which mixes concepts both from repeated games and the theory of incentives. The first chapter is a novel contribution to frequent monitoring in repeated games. The second one, studies for the first time, infinitely repeated games where the repetitions of the stage game are random. The last chapter, studies the provision of incentives in a principal-monitor-agent relation with exogenous learning, where the compensation is exogenously fixed. Each chapter can be considered independently of the rest.

Chapter I studies frequent monitoring in a simple infinitely repeated game with imperfect public information and discounting, where players observe the state of a continuous time Brownian process at moments in time of length Δ. It shows that a limit folk theorem can be achieved with imperfect public monitoring when players monitor each other at the highest frequency, i.e. $\Delta \to 0$. The approach proposed places distinct initial conditions on the process, which depend on the unknown action profile simultaneously and privately decided by the players at the beginning of each period of the game. The strong decreasing effect on the expected immediate gains from deviation when the interval between actions shrinks, and the associated increase precision of the public signals, make the result possible in the limit. The existence of a positive monotonic relation between payoffs and monitoring intensity is also found.

Chapter II studies repeated games where the time repetitions of the stage game are not known or controlled by the players. Many economic situations of interest where players repeatedly interact share this feature; players do not know exactly when the next time they will be called to play the stage game will be. We call this feature *random/stochastic monitoring*. We show that perfect random monitoring is always better then the classical perfect deterministic monitoring when the players' discount function is convex in the time domain.

Surprisingly, when the monitoring is imperfect but public, the result does not extend in the same absolute sense to all frequencies of play. The positive effect in the players' discounting is not always sufficient to compensate for a potential loss in the informational content of the public signals, due to the extra uncertainty on the repetitions of the stage game. However, we establish conditions under which random monitoring allows efficiency gains on the value of the best strongly symmetric equilibrium.

Chapter III studies a dynamic principal-monitor-agent relation where a strategic principal delegates the task of monitoring the effort of a strategic agent to a third party. The latter we call the monitor, whose type is initially unknown. Through repeated interaction the agent might learn his type. We show that this process damages the principal's payoffs. Compensation is assumed exogenous, limiting to a great extent the provision of incentives. We go around this difficulty by introducing costly replacement strategies, i.e. the principal replaces the monitor, thus disrupting the agent's learning. We found that even when replacement costs are null, if the revealed monitor is strictly preferred by both parties, there is a loss in efficiency due to the impossibility of benefitting from it. Nonetheless, these strategies can partially recover the principal's losses. Additionally, we establish upper and lower bounds on the payoffs that the principal and the agent can achieve. Finally we characterize the equilibrium strategies under public and private monitoring (with communication) for different cost and impatience levels.

La Tesis se compone de tres ensayos en Teoría Económica. Los dos primeros se enmarcan dentro de la Teoría de Juegos Repetidos, mientras que el tercero mezcla conceptos de esta teoría con la Teoría de Incentivos. El primer ensayo es una contribución novedosa a la literatura de monitorización periódica (frequent monitoring) en juegos repetidos. En el segundo se estudia, por primera vez en la literatura, juegos infinitamente repetidos en los que el intervalo entre dos repeticiones sucesivas del juego de etapa es aleatorio. El último ensayo analiza el problema de la provisión de incentivos en el marco de una relación principal-agente-monitor con aprendizaje exógeno, donde la compensación que recibe el agente por parte del principal se fija también de manera exógena. Cada uno de los ensayos es independiente del resto.

El Capítulo I plantea un modelo de monitorización periódica en un juego repetido infinito con descuento donde la información es pública pero imperfecta. La información se modela mediante un proceso Browniano en tiempo continuo. Los jugadores acceden a esta información en instantes discretos de tiempo, con frecuencia Δ. El valor inicial del proceso depende del perfil de estrategias que de forma simultánea eligen los jugadores en el momento inicial y, por tanto, es desconocido. Los jugadores penalizan las desviaciones del equilibrio del oponente. El resultado fundamental en este capítulo es un Folk-Theorem con información pública imperfecta cuando la frecuencia de monitorización tiende a cero, $\Delta \to 0$. Este resultado es posible debido a que, a medida que se reduce el período de supervisión, se mejora la precisión de las señales públicas, por lo que el jugador puede penalizar eficazmente las desviaciones de su oponente, reduciendo drásticamente los beneficios esperados inmediatos a una desviación. También se prueba que existe una relación constante positiva entre los pagos esperados y la frecuencia de supervisión.

El Capítulo II contiene una extensión del Capítulo I, pues se supone que los instantes de tiempo en que se producen las repeticiones del juego de etapa no son conocidos, sino que vienen determinados por algún proceso aleatorio. Denominaremos a esta situación monitorización estocástica. En muchas situaciones de interés económico donde los jugadores interactúan de forma repetida en el tiempo, los jugadores no saben exactamente cuándo es la próxima vez que serán llamados a jugar el juego de etapa. Por citar sólo algunos; dos empresas coludidas cambian sus estrategias dependiendo de los precios observados, que poden oscilar de una forma estocástica. Similarmente un superior puede monitorizar un empleado en momentos de tiempo de una forma imprevisible para esto.

El resultado principal en este capítulo es que si la función de descuento de los jugadores es convexa respecto al tiempo, entonces los pagos obtenidos bajo monitorización estocástica perfecta son mejores que los obtenidos en el caso clásico determinista. Por el contrario, cuando la monitorización es imperfecta el resultado no se extiende pata todas las frecuencias Δ. En este caso, el efecto positivo debido a la convexidad en la función de descuento no siempre es suficiente para compensar una eventual pérdida en la eficacia del contenido informativo de las señales públicas, debida a la incertidumbre en las repeticiones del juego de etapa. Es posible, sin embargo, aislar condiciones bajo las cuales la monitorización estocástica permite una mayor eficiencia en el valor del mejor equilibrio fuertemente simétrico.

En el Capítulo III se estudia una interacción dinámica principal-agente-monitor, en la que un principal delega la tarea de supervisar el esfuerzo de un agente a un tercero, que denominamos monitor. Tanto el principal como el agente se comportan estratégicamente, pero no así el monitor, quien puede presentar varios tipos y cuyo tipo inicial se desconoce. A través de la interacción repetida el agente puede aprender el tipo del monitor. Este proceso de aprendizaje reduce los pagos del principal. El hecho de que la indemnización se supone exógena, limita en gran medida el diseño de incentivos que eviten este problema. En lugar de utilizar este mecanismo, el principal puede reemplazar al monitor para impedir que el agente aprenda su tipo. Las estrategias de reemplazo tienen asociado un coste. Se demuestra que, incluso cuando el coste de reemplazo es nulo, si el monitor se revela estrictamente preferido por ambas partes, entonces hay una pérdida de eficacia debido a la imposibilidad de beneficiarse de esta preferencia conjunta. Sin embargo, el uso de estrategias de reemplazamiento puede recuperar en parte las pérdidas del principal. Además, establecemos límites superior e inferior de los pagos que el principal y el agente pueden obtener. Finalmente, se caracterizan las estrategias de equilibrio bajo control público y privado (con comunicación) para diferentes costes y niveles de impaciencia.

CHAPTER I

FREQUENT MONITORING IN REPEATED GAMES UNDER BROWNIAN UNCERTAINTY

1.1 Introduction

In the general repeated games theory, it is common to assume that the period length between each repetition of the stage game is of fixed length. When monitoring is perfect, letting the discount factor $\delta \to 1$ either by making the players more patient (a decrease in r) or by shrinking the period length between actions (a decrease in Δ) are equivalent exercises. The former approach has been preferred to prove many folk theorems and to show the existence of efficient equilibria.

When monitoring is imperfect, taking $r \to 0$ or $\Delta \to 0$ leads to different results. The reason is that variations in r and Δ have a different impact on the distribution of the signals. By making players increasingly patient through $r \to 0$, Fudenberg, Levine and Maskin (1994) were able to prove a folk theorem.[1]

The pioneer work in frequent monitoring is due to Abreu, Milgrom and Pearce (1991), in a setting where the public signals are modelled with a Poisson process. They have shown that different but inefficient results arise depending on whether $\delta \to 1$ is due to $r \to 0$ or to $\Delta \to 0$. In the latter case, they found that *strongly symmetric equilibrium* (SSE henceforth) payoffs above the static Nash, but not efficient, can be sustained when the jumps in the process represent "bad news", which are more likely to occur when some player has deviated.[2]

[1] Their stronger informational assumption is called *pairwise identifiability*, meaning that a deviation from a given player impact on the distribution of the public signals differently than any deviation from any other player. Incentives are provided through transfers of value between the players.

[2] In an infinitesimal time interval, the absence of realizations of the Poisson process is infinitely more likely than the occurrence of a realization. For that reason, the same result does not extend when the information arrivals represent "good news".

Recent work by Sannikov (2007)[3] and Faingold and Sannikov (2007) on repeated games modeled directly in continuous time, has renewed the interest in frequent monitoring. The latter work, in the part which is relevant here, reports a degeneracy of the set of SSE payoffs in a game where a known normal type long-run player faces a sequence of short run players. By degeneracy they mean that payoffs outside the convex hull of the Nash equilibria payoffs set cannot by sustained in continuous time, where the noisy public information is modeled through a Brownian process. More in the spirit of the present paper, i.e. by studying the limit of the discrete time games, Fudenberg and Levine (2007) and Sannikov and Skrzypacz (2007) report the same degeneracy result. These results came as a surprise, since Brownian motion is an infinitesimal variation process and we would expect payoffs at least above the static Nash payoffs, due to an increased precision of the signals.

This paper explores frequent monitoring in a partnership game with imperfect public monitoring and discounting.[4] It analyses the limit of the sequence of the discrete time games indexed by Δ. The public signal is the observed state of an arithmetic Brownian motion (ABM henceforth) process, in intervals of length Δ. The observation is compared against a previously chosen threshold, based on this decision rule; players adjust their actions for the following period.

A great deal of attention is given to SSE payoffs, not only because of their simplicity but also because with two-sided imperfect public monitoring, the pairwise identifiability assumptions typically fail, limiting to a great extent the provision of incentives. Destruction of value through punishments is the only way to provide incentives. Nonetheless, we show that the value of the best SSE payoff improves monotonically when the monitoring intensity increases. We also provide a full characterization of the optimal decision rule for different values of Δ.

Finally, we show that in the limit a folk theorem can be obtained, independently on how

[3] This paper provides a novel and elegant characterization of the set of *perfect public equilibria* payoffs using continuous time methods.

[4] Essentially, the information structure is similar to the one of Radner, Myerson and Maskin (1986). Other, classical situations involving imperfect public monitoring are Green and Porter (1984) and Porter (1983) where the market price is an imperfect signal of the quantities supplied by firms. See Fudenberg and Tirole (1991) and Mailath and Samuelson (2006) for complete surveys on repeated games.

players discount the future and on the level of uncertainty.[5]

In the context of the existing literature, these are striking results. The proposed approach, places distinct initial conditions on the process, which depend on the unknown action profile privately and simultaneously decided by the players at the beginning of each period of the game. This modelling approach changes the results drastically.

This paper is the first to show efficient results when the time interval between observations is taken to the limit without placing assumptions on the volatility of the process. The result is possible because the information extracted from the public signals becomes increasingly precise about the players' actions, increasing the payoffs monotonically.[6] Moreover, the expected immediate gains associated with a deviation from the equilibrium path become less attractive, not only because the periods between the actions becomes shorter, but also because the expected number of periods during which a deviator can enjoy these gains decreases.

Related Literature - Sannikov and Skrzypacz (2007)[7] study how monitoring intensity affects the equilibrium payoffs of a repeated Cournot duopoly game. They report the impossibility of achieving payoffs higher than the static Nash payoffs when the public information arrives continuously and disturbed by a Brownian motion. Crucial for their result is the assumption that the public signal observed by the players, at moments in time $t = \Delta, 2\Delta, ...$, is the state of an ABM price process divided by the length of the time interval Δ. Such modelling of the observed public signal becomes extremely noisy when observed at high frequency, creating a degeneracy effect on the payoffs. The root of the problem lies on the

[5] By "in the limit" we mean the length of time interval $\Delta \to 0$, sometimes also referred to as the "highest monitoring intensity" or "continuous monitoring". During the paper we frequently mention "an increase in the monitoring intensity" or "an increase in the monitoring frequency"; they refer to a decrease in Δ.

[6] The monotonicity results share similarities with the ones obtained in Abreu, Milgrom and Pearce (1991). For the "bad news" case, the most efficient equilibrium (not fully efficient) is obtained in the limit, and there is also a monotonic improvement on the payoff with the monitoring intensity. Here the most efficient result is also obtained in the limit; it is however fully efficient. In the "good news" case, the degeneracy of the best SSE reported in Abreu, Milgrom and Pearce, is similar in shape to the results obtained by Fudenberg and Levine (2007) and Sannikov and Skrzypacz (2007), although clearly distinct in the mechanics behind it. This issue will be discussed in more detail below.

[7] See also, Sannikov and Skrzypacz (2009) where they bound the set of equilibrium payoffs by placing restrictions on how information from Brownian and Poisson components are used to provide incentives in the most efficient way.

fact that the accumulated Brownian increments in a given time interval Δ are of a higher order than the underlying time interval, making any inference about the drift of the process inefficient in the limit.

Fudenberg and Levine (2007)[8] study a repeated game with a finite action space where a long-run player of a known type faces a sequence of myopic short-run players. The public signal is the observed state of an ABM process, which can be influenced by the actions of the long-run player. The long-run player would like to sustain the equilibrium associated with a particular action, in which case the drift of the process decreases with Δ, but she may also deviate to get a larger expected short run payoff, in which case the drift of the process increase with Δ. The results are driven by the assumption that the initial conditions on the process are the same independently of the actions of the long-run player. When Δ becomes small the distribution of the public signals cannot provide reliable information about the long-run player actions, creating the degenerating effect.[9]

However, they show that if a deviation by the long-run player increases the volatility of the process; equilibria can be achieved that are arbitrary close to efficiency. When deviating, the increase on the noise of the signals favours the provision of incentives, because inference becomes more precise, the decision rule relaxes and mistaken punishments vanish in the limit. The limit result obtained by Fudenberg and Levine is similar to the one presented in the present paper. However, the mechanics that lead to asymptotic perfect monitoring are different.

The same efficient result does not generalize when a deviation decrease the uncertainty parameter, as shown in Fudenberg and Levine (2007). The reason is that to detect a downward swift in the volatility and to keep the incentives, players incur too often in wrong punishments.

[8] See also, Fudenberg and Levine (2009) where they consider different ways of passing to the continuous time limit, i.e. binomial and trinomial approximations of the Brownian paths, linking monitoring intensity with event frequency, by taking the limit of the former. Such construction can also be applied in the context of the present paper, for example, by considering two binomial trees that start in different points associated with the different initial profiles of actions. The trees intercept in the second row or after.

[9] It can be shown that there is some equivalence between the Fudenberg and Levine (2007) and Sannikov and Skrzypacz (2007) approaches.

It is worth to notice, that while the variance parameter can be consistently estimated from the path of the process $(y_s, s \in (t, t+\varepsilon])$, for a small but measurable ε, the same does not happen with the drift of the process.[10] Even if players are able to observe the full path of the process realized from time t to time $t+\Delta$, a relatively large Δ is needed for the actions of the players to be statistically distinguishable. The first observation is on the basis of Fudenberg and Levine's (2007) efficient result. The second explains why equilibrium payoffs above the static Nash are not possible, when Δ decreases and players make inference over the drift of the process, as reported in Fudenberg and Levine (2007) and Sannikov and Skrzypacz (2007).

Discussion - The present paper models the public signal observed by the players at moments in time $t = \Delta, 2\Delta, ...$, differently. When each player privately selects her action, the initial condition of the process reflects the aggregate of these individual decisions. Two different actions of player i have associated different initial conditions that can be statistically distinguished from each other for high monitoring frequencies. There is a measurable distance between an initial condition associated with mutual cooperation and an initial condition associated with a profile of actions where a unilateral deviation has occurred. When switching from cooperation to defection, player i causes a movement in the process similar to a jump. When away from the limit, such jump might be hard to separate from the aggregate of infinitesimal realizations, but in the limit such movement is almost surely caused by deviating behaviour. This is even more evident when the action space is discrete, but is also true with a continuum of actions.

It is important to stress that the modelling proposed in this paper, and the Fudenberg and Levine (2007, 2009) and Sannikov and Skrzypacz (2007, 2009), are not competing approaches, but rather they complement each other. These approaches are conceptually correct, however, the degeneracy effect reported when player control the drift of the process

[10] See Prakasa Rao (1999) for a formal treatment of the statistical methods and the information that can be extracted from statistical inference for diffusion processes.

might be counterintuitive in some situations. The present paper presents a modeling approach to frequent monitoring that attempts to rationalize the intuition that more monitoring cannot harm the monitor side and has positive effects in terms of incentives for cooperation and higher effort.[11] In particular, we are thinking in problems where the informativeness of the public signals increases with the monitoring intensity. We want to develop a theory where imperfect public monitoring, in the limit, is equivalent to perfect monitoring.

Dickinson and Villeval (2008), show empirically that monitoring has a positive effect on individuals effort. However, there is also evidence of a Frey's (1993) crowding-out effect, which can lead to an equilibria degeneracy in the limit.[12] This effect is based on behavioral aspects, that are not considered in the present paper, neither on the existing literature on frequent monitoring, which rather focus on the informativeness of the signals with respect to the monitoring intensity.

Consider the following example, in line with the present paper. In an infinitely repeated Cournot duopoly game, firms supply choices are private information, but the market price is publicly observed.[13] If one of the firms deviates, by taking a decision (in the beginning of a given period) to increasing its own production flow, then the noisy market price should adjust instantaneously to this new aggregate output flow. When the periods between observations of the market price are large, and only the current market price is observed, such deviation is likely to pass undetected due to the aggregate of exogenous noisy events. However, when the current market price is observed at a high frequency the sum of exogenous noisy perturbations on the prices becomes negligible. The effect of an increase in the supply is almost surely detected.

[11] See, for example, the seminal paper of Alchian and Demsetz (1972) for an early defence of the disciplinary effect of monitoring. See also, Dickinson and Villeval (2008) and the references there in.

[12] Frey's idea, is that the monitoring crowding-out effect should be stronger in interpersonal relations. On the other hand, the disciplinary effect should dominate in abstract relations (more distant), where psychological issues and intrinsic motivation aspects are absent.

[13] In this paper, we show our results in a simpler game, because we want to focus on the informational issues without adding extra complexities. The Cournot game will be discussed in Section 1.6.

Fudenberg and Levine (2007, 2009) and Sannikov and Skrzypacz (2007, 2009) approaches, are suitable to model situations where reliable information needs time to build-in, in such cases an early or frequent access to this "not ready" information generates perverse effects. These papers have in common the fact that the informativeness of the public signals increases with the lack of monitoring. Clearly, such cannot fit in many economic problems of interest.

In this sense, the present paper fills a gap in the existing literature, by enlarging the spectrum of economic problems that can be studied using the theory of frequent monitoring.[14]

The paper is organized as follows. Section 1.2 presents the repeated game model and the public information producing process. Section 1.3 explains in detail the approach of this paper, in particular the connection between the initial conditions of the process and the associated distributions. Section 1.4 computes the bounds on the set of SSE payoffs and characterizes the optimal decision rule for varying Δ. Section 1.5 focuses on the limit case and presents the main results of this paper. Section 1.6 discusses extensions to the continuous time case and to games with a continuous action space. Section 1.7 concludes.

1.2 The Repeated Game Model

We explore frequent monitoring in a simple partnership game with two long-run players. The history of the game is the following. At moments in time $t = 0, \Delta, 2\Delta, ...$, players can choose from two different effort levels $a_{it} = 1$ or $a_{it} = 0$. In the former case, player i is providing effort E, in the latter case she is shirking S. More formally, let $A_{it} = \{0, 1\}$ denote player i's $\in N = \{1, 2\}$ non-empty and compact action space with generic element a_{it} representing an action, and denote $A_t = A_{1t} \times A_{2t}$ as the set of action profiles endowed with the product topology of the individual action spaces, with generic element $a_t = (a_{1t}, a_{2t})$ denoting a profile of actions.

[14] In fact, recent development on the theory frequent monitoring has allowed to study interesting departures from the canonical repeated game framework. For example, Fudenberg and Olszewski (2008) study the limit of an infinite horizon finitely repeated game with random asynchronous monitoring, while Osório-Costa (2008) study infinitely repeated games, where the repetitions of the stage game are not deterministic.

Independently of their private effort decisions, players at moments in time $t = \Delta, 2\Delta, ...,$ observe the realized total output $y_{t+\Delta}$ generated during these time intervals of length Δ. Given a profile of actions a_t chosen at time t, the public signal observed at time $t + \Delta$ (the total output),[15] is driven by the following *arithmetic Brownian motion* (ABM henceforth) process,[16]

$$y_{t+\Delta} = y_t + \sigma \int_t^{t+\Delta} dZ_s, \text{ with } Z_t = 0 \text{ and } t = 0, \Delta, 2\Delta, ..., \tag{1}$$

where $y_t \equiv 2\pi' (a_{1t} + a_{2t})$ is the initial condition of the process at a given time t, a function of the unknown profile of actions. The parameter σ, gives a measure of the volatility or noise of the process and π' is a productivity measure.

The uncertainty is generated by the standard Brownian motion $\{Z_s; s \geq 0\}$. Notice, that information about the evolution of the total output is produced continuously, in every infinitesimal instant of time a new realization of the process is available. The monitoring frequency Δ and the frequency of signals are different; in some sense our goal is to equate the monitoring frequency to the signal frequency.

All the relevant information about players' actions is contained in the initial condition of the process at each moment in time $t = 0, \Delta, 2\Delta, ...$ Since players cannot revise their actions during the interval of length Δ, we removed the drift of the process (1). In this way, we also eliminate any trend in the process, which in our setting is irrelevant for the study of the problem.[17]

Notice, that the public process is a martingale with respect to some filtration, i.e. $E(y_{t+\Delta} | y_t, \Delta \geq 0) = y_t$. The transition density of the process places equal mass above and below the initial condition, i.e. above and below its mean.[18]

[15] We could have also considered the possibility that in the end of each period of length Δ, players observe the full path of the process $\{y_s, s \in (t, t + \Delta]\}$ realized from t to $t + \Delta$. This case provides more information to the monitor. It can be shown, that when compared with the case where only the state of the process is observed, a lower threshold is required to sustain a particular equilibrium. Consequently, it always has larger associated payoffs. Note, however, that in the limit both cases are equivalent.

[16] The ABM assumption is made for simplicity. All results are valid for the geometric Brownian motion and Ornstein-Uhlenbeck process.

[17] With different initial conditions, the process (1) with drift leads to the same results. In that case, the threshold would be some function preserving the distance to the initial condition. For example, for the ABM process with drift, the threshold would be the function $b + \mu(y_t, t)\, t$, where b is the threshold associated with an ABM process without drift and $\mu(y_t, t)\, t$ is the drift of the process. Both approaches lead to the same distribution of the public signals.

[18] Examples of other processes with the proposed specifications are, $\mu(y_t, t) = 0$ for the geometric Brownian

Let, player i's $\in \{1,2\}$ realized payoff (ex-post) from the partnership be given by,

$$r_i\left(a_{it}, y_{t+\Delta}\right) \equiv y_{t+\Delta}/2 - \left(2\pi' - \pi\right) a_{it},$$

where $\pi' > \pi > 0$ and $y_{t+\Delta}$ is the realized output from the partnership, which is divided equally between the players independently of the effort that they have supplied.[19] The second term on the *right-hand side* (RHS Henceforth) denotes the cost of providing effort for player i. The expected payoff (ex-ante) of player $i \in \{1,2\}$ from the partnership is then,

$$\pi_i\left(a_t\right) \equiv E\left(r_i\left(a_{it}, y_{t+\Delta}\right)|y_t\right) = \pi'\left(a_{1t} + a_{2t}\right) - \left(2\pi' - \pi\right) a_{it}.$$

Under the expected utility hypothesis, this is the relevant expression for studying the problem.

In resume, at each moment in time $t = 0, \Delta, 2\Delta, ...$, players repeatedly play the stage game,

	$E \Leftrightarrow 1$	$S \Leftrightarrow 0$
$E \Leftrightarrow 1$	π, π	$-\left(\pi' - \pi\right), \pi'$
$S \Leftrightarrow 0$	$\pi', -\left(\pi' - \pi\right)$	$0, 0$

Shirking is a dominate strategy for both players. The minimax value of the game coincides with the stage game Nash payoffs and equals 0 for both players. For convenience we assume $2\pi > \pi'$. This game has the same structure as a prisoners' dilemma, and can be treated as such.

A great deal of attention will be given to the SSE payoffs. A strongly symmetric public strategies, is when after every public history the same action is chosen by both players.[20] More generally, a strategy is public if at any moment t it depends only on the public histories and not on player i's private history. Given a public history, a profile of public strategies

motion, and $\mu\left(y_t, t\right) = \rho\left(y_0 - y_t\right)$ for the Ornstein-Uhlenbeck process.

[19] We assume that public signal is not only an action dependent function but also represents the evolution of the aggregate output of the partnership. Other formulations of the public signal could have also been considered, provided that they would depend on both players' actions. It is also important that $r_i(.)$ does not depend on a_{-i} explicitly.

[20] A public history, a time t, is a sequence of realizations of the observed state of the process, denoted by $h^t = \left(y_{t-\Delta}, y_{t-2\Delta}, ..., y_0\right) \in Y^t$, with $h^0 = Y^0 \equiv \varnothing$. The sequence of player i's private effort choices, is player i's private history.

that induces a Nash equilibrium on the continuation game from time t on, is called a *perfect public equilibrium* (PPE henceforth). Moreover, if the other player $-i$ is playing a public strategy, player i's best reply can only be a public strategy.

Players discount the futures according to a common discount factor, assuming exponential discounting $\delta \equiv e^{-r\Delta}$, where r is the discount rate.

1.3 The Initial Conditions and the Distribution of Public Signals

As discussed in the introductory section, the proposed approach places distinct initial conditions on the process, which depend on the unknown action profile simultaneously and privately decided by the players in the beginning of each period of the game. Since this is a critical issue, in this section we examine in more detail the monitoring technology employed in this paper.

To keep the notation simple, from now on we drop the t index, and denote with a Δ when it refers to an end of period object, and without index when referring to the beginning of the period.

The signal space is continuous; players use a threshold b decision rule to distinguish realizations suggesting cooperation from realizations suggesting defection.[21] The cut-off value creates a partition of the signal space; in signals suggesting cooperative behaviour $\{y_\Delta > b\}$, which we call "good" signals, and signals suggesting defective behaviour $\{y_\Delta \leq b\}$, called "bad" signals.

In general, for a given initial condition $y \equiv 2\pi' (a_1 + a_2)$, the probability that the state of the public process (1) appears below b in the end of the period of length Δ is

$$\Pr (y_\Delta \leq b) = \Phi \left(\frac{b - 2\pi' (a_1 + a_2)}{\sigma\sqrt{\Delta}} \right),$$

where $\Phi (.)$ is the standard zero mean and unit variance Gaussian distribution.

We assume that players know the value of parameter σ and the type of uncertainty

[21] See, for example Sannikov and Skrzypacz (2007), where they show that a threshold decision rule is the best test to detect unilateral deviations. Later, we will focus on the optimal threshold value for varying parameterization of Δ. For now we contend with an arbitrary threshold, and we will abstain from referring its dependence on Δ as well as other parameters of the model.

they are facing.[22] This way they can compute the above probability and the impact of a deviation on the distribution of the public signals.

Depending on the unknown profile of actions that arise from each player's private effort decision $a_{i\in N}$, we have different initial conditions for the public process. In the partnership game there are four possible profiles of actions: the strongly symmetric profile $a \equiv (1,1)$, the asymmetric profiles $a' \equiv (1,0) \equiv (0,1)$ and the Nash profile $a^N \equiv (0,0)$, which is trivially self-enforceable.[23]

(i) When we want to enforce the SSE profile $a \equiv (1,1)$ we have, the initial condition or ex-ante output, $y = 4\pi'$. Then probability of punishment is given by

$$F^{EE} \equiv F^{1,1}(b,\Delta) \equiv \Phi\left(\frac{b-4\pi'}{\sigma\sqrt{\Delta}}\right).$$

This is type I error probability; even though no player has deviated, punishment will be exerted when a low realization of the process is observed.[24] We want to protect this profile against a potential deviation, $a' \equiv (0,1) \equiv (1,0)$. In this case, the profile a' has associated an initial condition $y = 2\pi'$, and a probability of punishment given by

$$F^{ES} \equiv F^{1,0}(b,\Delta) = F^{0,1}(b,\Delta) \equiv \Phi\left(\frac{b-2\pi'}{\sigma\sqrt{\Delta}}\right).$$

Analogously a type II error is given by $1 - F^{ES}$.

(ii) The most asymmetric equilibria payoffs for player i, require a initial profile of play $a' \equiv (1,0) \equiv (0,1)$, we denote the probability of mistaken punishment as

$$G^{ES} \equiv G^{1,0}(\underline{b},\Delta) = G^{0,1}(\underline{b},\Delta) \equiv \Phi\left(\frac{\underline{b}-2\pi'}{\sigma\sqrt{\Delta}}\right).$$

Notice, that when considering to deviate, player i would do it in the first period, because it is when the worst profile is due. The deviation by the player i that is providing high effort,

[22] The assumption that players know the variance of the process is not as strong as it seems. According to the discussion in the introductory section, since the focus is on the limit case and this is equivalent to full path observation, players can estimate the true value of σ consistently in a very small but measurable time interval ε.

[23] Fudenberg, Levine and Maskin (1994) pairwise full-rank condition typically fails for strongly symmetric equilibria. There is no loss in generality when placing no distinction between the profiles $(1,0)$ and $(0,1)$, and denoting them as a'. Also note that to keep the notation standard, until now, a has denoted a general action profile. With a slight abuse of notation, a now denotes the strongly symmetric effort profile.

[24] We refer to a type I error as the event of punishing a non-deviator, and a type II error as the event of not punishing a deviator.

leads to the profile $a^N \equiv (0,0)$. The initial condition is then $y = 0$, and the probability of punishment given by

$$G^{SS} \equiv G^{0,0}\left(\underline{b}, \Delta\right) \equiv \Phi\left(\frac{\underline{b} - 0}{\sigma\sqrt{\Delta}}\right).$$

Notice that, depending on the initial profile we want to sustain, different thresholds have to be computed. For that reason we distinguish between b and $\underline{b}$. Then, F^{ES} and G^{ES} do not have the same value, since they have associated different decision rules. We employ the distributions G^{ES} and G^{SS}, when our goal is to enforce a path that starts with the profile $(1,0)$, while F^{EE} and F^{ES} are employed when enforcing $(1,1)$. The former set of distributions, will be used only in Section 1.5, to prove the main result of this a paper.

Observe that the actions decided independently at the beginning of the period by each of the players are clearly printed in the initial conditions of the process, as they should. At the end of the period of length Δ, the initial condition associated with a given profile of actions appears disturbed by some noise. Lowering Δ, we are decreasing the uncountable number of infinitesimal contributions to the noise, leading to a more likely observation of the process around its initial condition. An observation of the process far from the initial condition is then a sign of a deviation.

If we want to generalize this methodology to other more complex environments, this correspondence between actions taken and the associated expected public signal observed in the end of the period has always to be present.

Active Coordination

When the game is a repeated partnership, it is natural to assume that the process is reset at the end of each period, after players have observed and split the realized output, starting again in the point associated with the new action profile decided simultaneously by both players. However, the same assumption does not fit in some other contexts. For example in the repeated Cournot game, it is not reasonable to assume that the process is reset every period, since it represents the market price. In this case, the new equilibrium actions have to be adjusted to take into account the observed state of the process.[25] Although different, these two possibilities can be studied within the same methodology. To illustrate this point consider the following example.

In the Cournot duopoly game the public process observed at time Δ is the market price, $P_\Delta = P + \sigma \int_0^\Delta dZ_t$, where $P = (\alpha - q_1 - q_2)$. Suppose that at time 0 the game starts with both players supplying the monopoly quantities, i.e. $q_i = \alpha/4$.[26] The expected market price (and initial condition) is then $P = \alpha/2$ and the threshold value is some $b < \alpha/2$.

Let the end the period realized market price be, $P_\Delta = P + \sigma z_\Delta$. If $P_\Delta \leq b$, players play their punishment actions, with respect to the observed state of the process. For simplicity, consider the Nash punishment; in the following period each player will supply $q_{i\Delta} = (\alpha + \sigma z_\Delta)/3$.

On the other hand, if $P_\Delta > b$, for the next period, the 50/50 monopoly quantities will be chosen, taking into account the state of the process, hence, requiring each player to supply the quantities $q_{i\Delta} = (\alpha + \sigma z_\Delta)/4$ with a necessary adjustment in the decision rule for the next period, i.e. $b_\Delta = b + \sigma z_\Delta$.

We will call this type of tacit coordination in a dynamic setting *active coordination*, since it requires players to adjust their equilibrium path actions taking into account the observed state of the process. Players do not reset the value of the process, rather they reset the uncertainty.

1.4 The Best Strongly Symmetric Equilibria

This section furnishes the reader with a set of general results that are independent of the monitoring intensity, which are particularly useful for the following section when we focus on the limit case. It also presents a characterization of the optimal decision rule associated with the value of the best SSE of the infinitely repeated partnership game. In some occasions, the results presented are standard in repeated games; a brief description will be presented,

[25] The martingale condition guarantees that such a procedure does not change the properties of the distribution of the public signals.

[26] In the Cournot duopoly game under imperfect public information, it is never optimal to produce exactly the monopoly quantities, but rather an amount slightly larger (except in the limit). This issue and others are discussed in more detail in section 1.6.2. Take this example as an illustration.

sufficient to keep the exposition self-contained.

Since the game is symmetric there is no loss in generality when studying a single player incentives, hence we remove players indexes.

The equilibrium profile we want to sustain in a SSE is $a \equiv (1,1)$. A profile where a single player deviates is denoted as a', and the Nash profile denoted as a^N. These profiles have associated the stage game payoffs, π, π' and 0, respectively. See the expected payoffs matrix in section 1.2. The public information is produced continuously and is generated by the ABM process given in (1), with initial conditions depending on the unknown action profile.

Since the public information process can virtually return any value in $\mathbb{R}$, we apply the Abreu, Pearce and Staccetti (1986, 1990) bang-bang result to compute the best SSE payoff. Punishments are not executed in terms of a deterministic number of time periods as suggested by Porter (1983) and Green and Porter (1984) but rather in a probabilistic sense. As we shall see, since the distribution of the public signals is not convex, see Porter (1983), optimality requires an infinite punishment length and both approaches do not differ. In some other problems, in particular when the minmax value differs from the Nash value, the differences might be substantial.

The expected continuation value of the game is some convex combination between the expected normalized payoff $\overline{v}$, when play starts with the observation of a good signal, and the expected normalized payoff $\underline{v}$, when play starts with an observation of a bad signal.[27]

The Abreu, Pearce and Staccetti result allows us to write players' problem in a very tractable way, and to use recursive dynamic programming methods to search for the expressions for the values of $\overline{v}$ and $\underline{v}$ that are exclusively represented as functions of the parameters of the model.

To find the expression that characterizes the best SSE payoff, we need to solve the

[27] Since later we will allow players to correlate their actions on some public signal, $\overline{v}$ and $\underline{v}$ are extreme points of the set of SSE payoffs. This set is the collection of all payoffs that can be achieved with strongly symmetric public strategies, i.e. when both players choose the same action after every public history. See also the discussion after Lemma 1.

following dynamic programing problem, which by symmetry is the same for both players:

$$\overline{v} = (1-\delta)\,\pi + \delta\left[\left(1-F^{EE}\right)\overline{v} + F^{EE}\underline{v}\right], \tag{2}$$

$$\overline{v} \geq (1-\delta)\,\pi' + \delta\left[\left(1-F^{ES}\right)\overline{v} + F^{ES}\underline{v}\right], \tag{3}$$

$$\underline{v} = (1-\delta)\,0 + \delta\left[p\underline{v} + (1-p)\,\overline{v}\right], \tag{4}$$

$$p \in [0,1]\,. \tag{5}$$

Expression (2) is the value of the relation when both players always provide effort. While both players provide effort, each receives the immediate discounted normalized expected payoff associated with mutual effort, as well as a discounted expectation over the expected values $\overline{v}$ and $\underline{v}$, associated with the two types of signals that might be observed. The first constraint (3) is a enforceability condition. It has a simple interpretation; the expected value of the game associated with mutual effort has to be at least as good as the expected value of the game associated with a potential unilateral deviation, even if this deviation just last one period. When satisfied, (2) and (3) enforce the profile $(1,1)$.

Expression (4) is the value of the punishment phase, where p is the probability with which the relation remains in this stage. A value $p=1$ means perpetual punishment and a $p=0$, requires a single punishment period. Since $(0,0)$ is a Nash equilibrium, punishment is trivially enforced.

Condition (5), requires p to be a probability. Consequently, we must have $\underline{v} \in [0, \delta\overline{v}] \subset [0, \overline{v})$.

Expression (4) can be solved for $\underline{v}$ to obtain

$$\underline{v} = \frac{\delta\,(1-p)\,\overline{v}}{1-\delta p}. \tag{6}$$

Plugging $\underline{v}$ into (2) and (3) and making the latter hold with equality. We obtain the enforceability condition

$$\frac{1}{\delta} - \frac{F^{ES}\pi - F^{EE}\pi'}{\pi' - \pi} \equiv P = p. \tag{7}$$

We must have P taking a value in the interval $[0,1]$ in order for (7) to bind. If there is no way to obtain a value of $P \leq 1$, no equilibria other than the infinite repetition of the static Nash can be sustained, i.e. $\overline{v} = 0$. When this turns out to be the case, in particular in the limit, we say that the set of SSE payoffs degenerates. When $P < 0$, we don't have such a problem, as we will see below, we can adjust the optimal decision rule in order to obtain a value of $P \in [0,1]$.

After replacing (6) and the enforceability condition (7) into (2), we can solve the latter for $\overline{v}$ to obtain the expression of the best SSE

$$\overline{v} = \pi - \frac{F^{EE}}{F^{ES} - F^{EE}} \left(\pi' - \pi\right). \tag{8}$$

Similarly, replacing (7) and (8) into (6), we obtain

$$\underline{v} = \overline{v} - \frac{(1-\delta)}{\delta} \frac{\pi' - \pi}{F^{ES} - F^{EE}}. \tag{9}$$

Expression (8) and (9) characterize the value of the upper and lower bounds on a set of SSE payoffs respectively.

The following result establishes conditions on the threshold b in order to obtain the largest value $\overline{v}$, constrained by the fact that (7) holds with equality and it is feasible.

Lemma 1 *Under (1), the strategy that achieves the best SSE payoff $\overline{v}^*$, requires perpetual punishment the first time the process is observed below $b^*(\Delta)$. Where $b^*(\Delta) \equiv b^* \leq b^p$ is called the optimal threshold and solves $\underline{v} = 0$, i.e.*

$$F^{ES}\pi - F^{EE}\pi' = (1-\delta)\left(\pi' - \pi\right)/\delta, \tag{10}$$

and

$$b^p = 3\pi' + \sigma^2 \Delta \ln\left(\pi/\pi'\right)/2\pi', \tag{11}$$

is an upper bound on the optimal decision rule.

The result tells us that, independently of Δ, among all the feasible punishment schemes, perpetual punishment is the optimal one, i.e. expression (8) returns the largest SSE payoff. This is the case, because the ABM is a Gaussian process and the distribution of the public signals is not convex in its entire domain.[28]

[28] For a discussion of this issue in the context of repeated games, see Porter (1983).

The equality (10) gives an implicit function to compute b^*. The value b^* establishes the right balance between gains and losses associated with right and wrong inference about players' actions, while keeping incentives satisfied. Moreover, if an optimal decision rule b^*, exist it must take a value smaller than b^p. Later, we will see that in this region b^* is the unique value that satisfies (10). When a solution b^* to (10) does not exists, no other cut-off decision rule can enforce the profile $(1,1)$.

The upper bound b^p imposes any particular restriction in the choice of b^*, other than guaranteeing uniqueness. In the interval $(-\infty, b^p]$, we have $\partial F^{ES}/\partial b \geq \partial F^{EE}/\partial b$, i.e. the likelihood ratio, is larger or equal to one. An alert of deviation is more likely to occur when somebody has deviated. The likelihood difference $F^{ES} - F^{EE}$ is non-negative and increases with b.

Notice that the two point set $\{0, \overline{v}^*\}$ associated with (10) in Lemma 1 is self-generating, since the continuation values π^N and $\overline{v}^*$ are elements of the set. However, if we allow for public correlation, we can convexify the set $\left\{\pi^N, \overline{v}^*\right\}$, and we can say that any payoff in the interval $\left[\pi^N, \overline{v}^*\right]$ can be sustained as a PPE of infinitely repeated partnership game.

It is worth noticing, that when we consider sets of the type $[\underline{v}, \overline{v}]$ with $\underline{v} > 0$, public correlation is required in order for the set to be self-generating. The reason is that the continuations associated with the value of the game that starts with effort, and the value of the game that starts in the punishment phase must be in $[\underline{v}, \overline{v}]$.

The important point here is that such a generalization allows the set of SSE to be an interval. Moreover, Proposition 7 of Section 1.5 will be shown to hold, not under optimal behaviour but for more general sets of the type $[\underline{v}, \overline{v}] \subseteq [0, \overline{v}^*]$.

However, not explicitly mentioned and in order to keep the notation simple, the solution b^* depends on all the parameters in the model, i.e. π', π, r, Δ and σ. In particular, since the values of F^{EE} and F^{ES} are endogenously determined, expression (8) necessarily depends on how much players discount the future, through b^*.

The function P is strictly convex in b with a minimum value at point b^p. When $P(b^p) < 1$, then there must be two threshold values, say b_1 and b_2, that satisfy $P(b_1) = P(b_2) =$

1. The decision cut-off value is not unique. Suppose, $b_1 \leq b_2$, with associated strongly symmetric payoffs $\overline{v}_1 \geq \overline{v}_2$ respectively. In this case, it is not admissible to choose a threshold other than the one associated with the larger strongly symmetric payoff, i.e. b_1 dominates b_2.

Definition 2 *Given a value P, we say that a threshold value b is admissible when it has associated the largest payoff $\overline{v}$. If in addition $P(b) \in [0,1]$ we say that such threshold is also feasible.*

Expression (11) establishes an upper bound on the optimal decision rule, but also guarantees that any threshold below b^p is admissible. However not all $b \leq b^p$ are feasible, in particular, if the unrestricted minimum value $P(b^p)$ is larger than one, we cannot enforce the profile $(1,1)$. The set of feasible and admissible thresholds is then defined as $\{b \leq b^p : P(b) \in [0,1]\} \subseteq [b^*, b^p]$.[29]

Even though that there might exist a continuum of feasible threshold values, since $\partial\overline{v}/\partial b < 0$, we are mainly interested in the optimal choice b^*. Depending on the parameters of the problem, for sufficiently large Δ the set of feasible threshold values might vanish. In pure strategies, the value of Δ such that, $0 = \max_b \underline{v}$ or equivalently $1 = \min_b P$, corresponds to the monitoring frequency after which the profile $(1,1)$ cannot be enforced any more. Denote this monitoring intensity by $\overline{\Delta}$.

Lemma 3 *The value $\overline{\Delta}$ is the solution for Δ of (10) at the point (11).*

Since Brownian signals are Gaussian, we cannot obtain $\overline{\Delta}$ in close form. The result tells us how to obtain this value implicitly.

It is worth noticing that $\underline{v}\left(b^*\left(\overline{\Delta}\right)\right) = \underline{v}\left(b^p\left(\overline{\Delta}\right)\right) = 0$, then $b^*\left(\overline{\Delta}\right) = b^p\left(\overline{\Delta}\right)$. At $\Delta = \overline{\Delta}$, if we take a threshold value $b \neq b^p\left(\overline{\Delta}\right) = b^*\left(\overline{\Delta}\right)$, then $\underline{v}\left(b \neq b^p\left(\overline{\Delta}\right) = b^*\left(\overline{\Delta}\right)\right) < 0$ or equivalently $P\left(b \neq b^p\left(\overline{\Delta}\right) = b^*\left(\overline{\Delta}\right)\right) > 1$. For $\Delta > \overline{\Delta}$ the set of feasible thresholds vanish, and no other equilibrium than the infinite repetition of the static Nash can be obtained.

[29] We could have defined admissibility in different terms. Notice that the function $\underline{v}$ is strictly concave in b, taking a unique unrestricted maximum value at some $b^v \in (-\infty, b^p]$, that is necessarily smaller than $\overline{v}$. When at the point b^v we have $\underline{v}(b^v) > 0$ there are two threshold that satisfy the optimality condition $\underline{v}(b_1) = \underline{v}(b_2) = 0$. The lower of these two thresholds is the admissible one. The set of feasible and admissible threshold would be $\{b \leq b^v : \underline{v}(b) \geq 0\} \subseteq \{b \leq b^p : P(b) \in [0,1]\}$.

The monitoring intensity $\overline{\Delta}$ has the following asymptotic properties; when σ or r goes to ∞ we have $\overline{\Delta} \to 0$, while if $\sigma \to 0$ we obtain $\overline{\Delta} \to \ln\left(\pi' / \left(\pi' - \pi\right)\right)/r$, and if $r \to 0$ we get $\overline{\Delta} \to \infty$. Consequently, providing that σ and r are bounded, the interval $\left(0, \overline{\Delta}\right)$ is always guaranteed to be nonempty, i.e. $\overline{\Delta} > 0$.

By Lemma 1, if a solution $b \leq b^p$ to (10) exists, it is an optimal threshold b^*. The following result establishes that such a solution, indeed exists, and it is unique and differentiable in the interval $\left(0, \overline{\Delta}\right)$. The value $\overline{\Delta}$ depends on all the parameters of the model, and it is the lowest observational frequency of the public process, that can support some nontrivial equilibrium in the infinitely repeated partnership.

Lemma 4 *For all $\Delta \in \left(0, \overline{\Delta}\right)$, with $\overline{\Delta} > 0$, a solution $b^*\left(\Delta\right) \in \left(-\infty, b^p\right)$ to (10) exists, is unique and differentiable.*

Intuitively, depending on the parameters of the model, $\overline{\Delta}$ might take a larger or a smaller value. For a given Δ, the threshold $b^*\left(\Delta\right)$ adjusts to keep (10) holding with equality. When $\Delta \in \left[\overline{\Delta}, \infty\right)$ we cannot guarantee existence and differentiability of the function $b^*\left(\Delta\right)$. Shirk becomes a dominate strategy for both players. While, if $\Delta \in \left(0, \overline{\Delta}\right)$ we can find a differentiable function $b^*\left(\Delta\right)$, for all Δ in this interval, which makes (14) hold.

It is worth noticing that when $r \to 0$, i.e. the players become very patient, the value $\overline{v}^* \to \pi$. The best SSE payoff is full efficient. This happen because the distribution of the public signals is Gaussian and has unbounded support. So as r decreases so does the optimal threshold relaxes, taking lower values. The punishment probability decreases. In the limit, $r \to 0$, leads to $b^* \to \infty$. Such result suggests a folk theorem in the limit as $r \to 0$. We do not elaborate more on this issue because the result depends on the unbounded support of the Gaussian distribution. The proof of such result involve a characterization of the rate at which incentives for low effort decrease, when $r \to 0$, and the rate at which $F^{EE} \to 0$.

The Optimal Threshold - Numerical

Lemma 1 shows that $\overline{v}^*$ is obtained using the most severe punishment. Such punishment has associated with it an optimal decision rule whose existence is guaranteed by Lemma

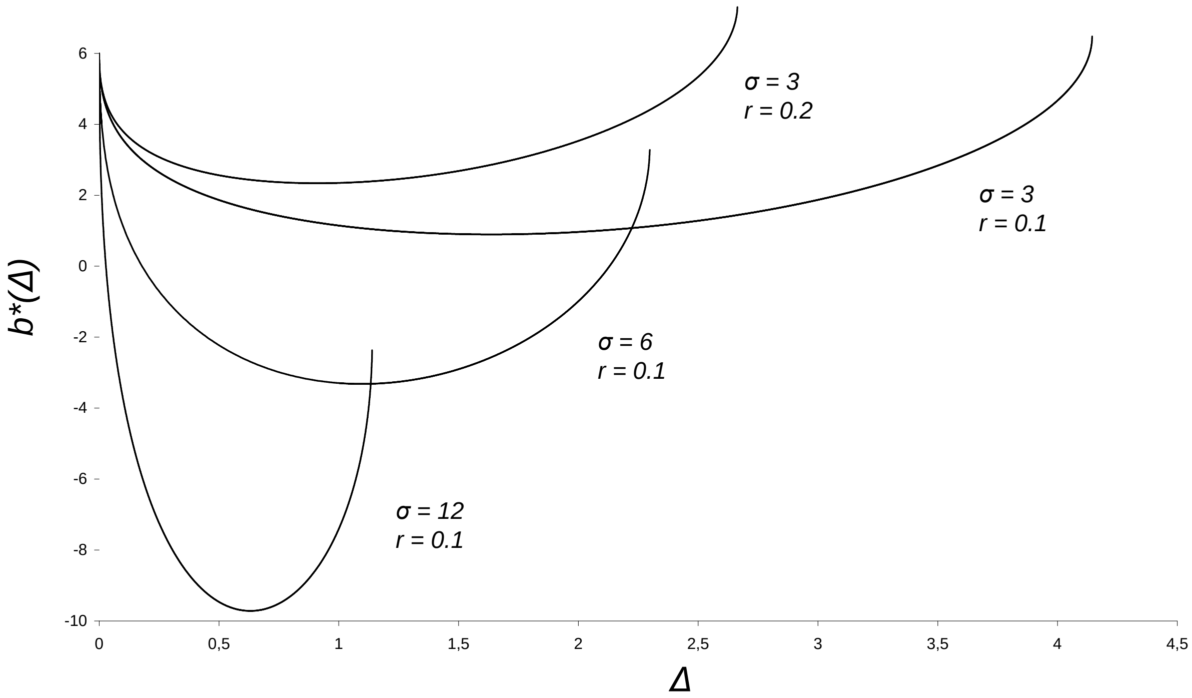

Figure 1: The optimal threshold as a function of Δ.

4. Now we will attempt to provide the intuition behind some properties of the resulting decision rule.

Figure 1, illustrate the optimal threshold value as a function of Δ, for the partnership game, and for different parameterization of σ and r, when $\pi' = 3$ and $\pi = 2$.

Notice that the value $\overline{\Delta}$ increases either when players get more patient or the public signal becomes less noisy, see Figure 1.

The more impatient the players are (larger r) the tighter the monitoring must be in order to create effort incentives.[30] This can be seen in Figure 1, when we increase r from 0.1 to 0.2, keeping σ fixed.

When we consider the noise parameter of the public process σ, it is not always true that large uncertainty leads to a lower threshold. For high monitoring intensities this is the case,

[30] By "tighter monitoring" as opposed to "relaxed monitoring", I mean a higher threshold value. The larger the threshold the more likely players are to detect deviations, but they are also more likely to make wrong judgements.

but when the monitoring frequency is low a tighter threshold might be required even if the uncertainty level is higher. The reason is that the incentives for deviation increase with Δ but also with noise of the signals. In Figure 1, when $\sigma = 6$ and Δ is around 2.3, we observe a larger threshold when $\sigma = 3$.

The strict convex shape of the threshold function with respect to $\Delta \in \left(0, \overline{\Delta}\right)$ is caused by two effects that operate in the same direction. As the monitoring intensity increases, i.e. Δ becomes small, the public signals become more informative about the other player private action. On same time, the expected immediate gains from deviating behaviour decrease. For infinitesimal Δ, the public signals get extremely informative and the expected immediate gains from deviation become negligible; for that reason the optimal threshold approaches $2\pi'$, the value to where the output would fall in case of perfect monitoring, this happens independently of the values that σ and r can take. Such a result is formally shown in Lemma 5 of Section 1.5.[31]

When the monitoring intensity decreases, the sum of infinitesimal variations of the process become more likely to generate a "bad signal". Wrong punishments in the equilibrium path are then more likely, for that reason the cut-off threshold relax, but in a decreasing way, because at the same time the expected immediate gains from a deviation become more attractive. At a certain point, after reaching its minimum value, the optimal threshold starts increasing at an increasing rate, creating the convex shape. This happens because the expected gains from deviation become increasingly attractive and the decision rule has to become tighter in order to keep players with incentives. Finally, for large values of $\Delta \geq \overline{\Delta}$, there is no threshold value that can sustain mutual effort; in pure strategies, shirk becomes dominant.

1.5 Monitoring Frequency and Limit Efficiency

In this section we look at the limit value of b^* and we will show that under Brownian uncertainty, monitoring intensity always has a positive effect on the payoffs. Finally, we

[31] It can also be shown that $\partial b^*/\partial\Delta \to -\infty$ when $\Delta \to 0$ and $\partial b^*/\partial\Delta \to \infty$ when $\Delta \to \overline{\Delta}$, however such results are not particularly relevant.

will present a Δ-limit folk theorem. These results are independent of how much players discount the future and of the uncertainty level.

1.5.1 The Limit Value of the Optimal Threshold

We start by studying the limit value of the optimal decision rule b^*. The question is; where the optimal threshold converges as monitoring become more and more frequent? Figure 1 above provides an illustration, the following result formalize it.

Lemma 5 *When $\Delta \to 0$ the optimal threshold $b^*(\Delta)$ converges to $2\pi'$, i.e. the expected signal associated with the deviation with less impact on the distribution of the public signals.*

Because different action profiles have associated different initial conditions, that are measurable separable from each other, in the limit the signals becomes perfectly informative about players actions. We have asymptotic perfect monitoring. Then $b^*(\Delta)$ must converge to the point where under perfect information, after a deviation, the realized deterministic output would be observed.

For the setting of our particular game, the result simply states that $b^* \to 2\pi'$, in the limit. However, the result can be more general; b^* must converge to the expected signal associated with the deviation with less impact on the distribution of the public signals. This is true even if the game has a continuous action space.

As mentioned before, the functional form of b^* cannot be obtained in close form, but clearly must depend on all the parameters of the model. For that reason, it is nontrivial to verify the rate at which b^* converges to $2\pi'$. A wide number of numerical simulations suggest that this convergence must occur at a rate lower than or equal to $\sqrt{\Delta}$. Suppose that the optimal threshold function has the following structure: $b^*(\Delta) = 2\pi' - \Delta^{\alpha} k(.)$ with $\alpha > 0$, where $k(.)$ is some function of the all parameters of the model and converges to some bounded value. While choosing $\alpha \geq 1/2$ we can show that efficient and feasible results hold in the limit. On the other hand, when choosing an α smaller than $1/2$ we can bound b^* from below, but the enforceability condition (10) fails. These results seem to suggest that $b^* \to 2\pi'$ at a rate of $\Delta^{0.49(9)}$. We do not develop this idea further, because of its complexity and because such result is of little relevance for the propose of the present paper. Nonetheless, the remark is left here.

1.5.2 Monotonicity of the Best SSE Payoff

A relevant questions, is how the value of $\overline{v}^*$ change with Δ? The following result establishes monotonicity between the value of the optimal upper bound on the set of SSE payoff and monitoring intensity. The result holds for any monitoring intensity $\Delta \in \left(0, \overline{\Delta}\right)$.

Proposition 6 *For $\Delta \in \left(0, \overline{\Delta}\right)$, the best SSE payoffs $\overline{v}^*$ increase monotonically with the monitoring intensity, i.e. with a decrease in Δ.*

In our setting, the monotonicity of the best SSE payoff to the monitoring intensity, suggests that more frequent the monitoring is larger are the payoffs. More monitoring is then preferred to less monitoring. The result is a consequence of the increased precision of the public signal when Δ becomes small. In a different context, Kandori (1992) shows a similar result, where an exogenous improvement in the precision of the public signals expands the set of PPE payoffs.

Figure 2 illustrate the strict monotonic improvement in the best SSE payoffs towards efficiency as the monitoring increases, in line with the statement of Proposition 6. Monotonicity is present independently on how players discount the future r, and the noisy parameter σ.

1.5.3 The Δ-Limit Folk Theorem

Figure 2 illustrates the value of the best SSE payoff $\overline{v}^*$, associated with the thresholds in Figure 1, converging to the full efficient outcome π. In this Section we are not only interested on the best SSE payoff $\overline{v}^*$, but we want to know to where the full set of equilibria converges in the limit as $\Delta \to 0$.

For simplicity, we focus on the case with unfavorable asymmetric payoffs for player 1, the other asymmetric case follow by symmetry.

According to Lemmas 1 and 4 of Section 1.4, we have the relation $\overline{v}^* > \underline{v}^* = 0$, for all $\Delta \in \left(0, \overline{\Delta}\right)$, with $\overline{\Delta}$ depending on the parameters of the model, and $\overline{v}^* = \underline{v}^* = 0$, for all $\Delta \in \left[\overline{\Delta}, \infty\right)$, i.e. the infinite repetition of the static Nash payoff.

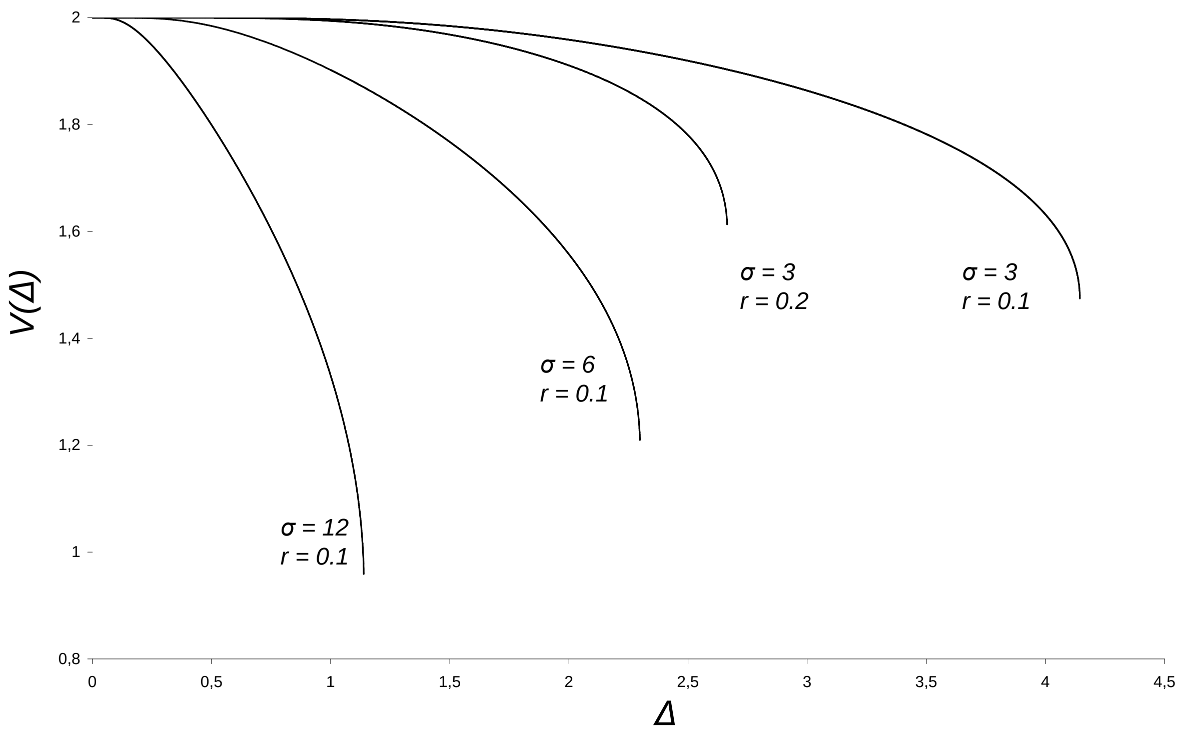

Figure 2: The best SSE payoff as a function of Δ.

Observe that a threshold that is slightly larger than b^*, but below b^p, still satisfies the enforceability conditions (10) with slack and at the expense of a lower SSE payoff. On the other hand values lower than b^* cannot enforce the profile $(1, 1)$. The threshold b^* gives us exactly the point that maximizes $\overline{v}$ according to Lemma 1. As we will see, limit efficiency does not require an optimal value for b, as in the limit, the same results hold with a larger threshold that is both feasible and admissible.

The punishment stage is not absorbing anymore; returning to the effort path is then possible. Consequently, we have the following relations, $\overline{v}^* \geq \overline{v} \geq \underline{v} > \underline{v}^* = 0$ with $b^* < b < b^p$ for all $\Delta \in \left(0, \overline{\Delta}\right)$. This is the strategy employed in the proof of the main result of this paper; if $\overline{v} \to \pi$ and $\overline{v} \leq \overline{v}^*$, then $\overline{v}^*$ must converge to π as well.

In the Lemmas 1, 3, 4 and 5, we developed a great knowledge about the threshold that sustains the best SSE payoff. About the optimal threshold associated with an asymmetric path that starts with the profile $(1, 0)$, the question is more delicate.[32] Nonetheless, following the discussion after Lemma 5, such threshold $\underline{b}^*$ must converge to 0, the initial condition associated with no effort from both players. However, there is no guarantee it that such convergence occurs from below, such would depend on the particular path in question. Additionally, any choice $\underline{b} \in (0, \pi')$ must be feasible and admissible, at least in the limit $\Delta \to 0$.

Recall, that F^{ES} and G^{ES} do not take the same value, since they have associated different decision rules. We use the distributions G^{ES} and G^{SS} when our goal is to enforce a path that starts with the profile $(1, 0)$.

In order to show that a folk theorem holds for the partnership game we apply the techniques of decomposition on tangent half spaces and the linear programming algorithm developed by Fudenberg and Levine (1994) and Fudenberg, Levine and Maskin (1994). Let $E^{set}(r, \Delta)$, denote the set of PPE payoffs that can be sustained for a given discount rate r and a monitoring intensity Δ, and denote $M^{set}(\Delta, r)$ the set of equilibrium payoffs that

[32] Notice that in general an asymmetric path, does not need to start with the profile $(1, 0)$, only the most asymmetric ones. However, such assumption is enough for our goal. It is also not necessarily true that the optimal moment for a deviation is the first period, but it is must be when the profile $(1, 0)$ is due.

result from such decomposition. This approach is preferred, since the set $E^{set}(r,\Delta)$ is typically hard to characterize. Clearly, since the set $M^{set}(\Delta,r)$ is a bound on the set of PPE, we have $E^{set}(r,\Delta) \subseteq M^{set}(\Delta,r)$.

It is worth to notice that in our setting, the set $M^{set}(\Delta,r)$, is not a limit set for $r \to 0$, since the distribution of the public signals depend on the optimal decision rule, which depends on r. A decrease in r expands the size of the bounding set $M^{set}(\Delta,r)$ as well as the set $E^{set}(r,\Delta)$.

Proposition 7 (The Δ-limit folk theorem for the partnership game) *Providing that the r and σ are bounded, every feasible and individual rational payoff of the infinitely repeated partnership game can be sustained as an equilibrium when $\Delta \to 0$.*

To obtain an efficient result in the limit we just need $F^{EE} \to 0$, $F^{ES} \to 1$, $G^{ES} \to 0$ and $G^{SS} \to 1$. When $\Delta \to 0$, any effect that r could have on players payoff is shadowed by the perfect informativeness of the public signals and the infinitesimal time period that a player can benefits from shirking behavior. In other words, under Brownian uncertainty when the monitoring intensity is taken to the limit, the public signals become perfectly informative about players' actions. Moreover, the effect of discounting r becomes irrelevant because any deviation is detected almost surely in the following infinitesimal time period.[33] In the limit the potential gains from deviating behaviour are negligible. Exception is when $r \to \infty$, in which case player do not discount.

We also need a bounding condition on the noise parameter σ.

The increased informativeness of the Brownian public signals for high monitoring intensity is the key aspect. It is due to the measurable distance between different initial conditions of the process.[34] Such a distance in a process of infinitesimal variation plays a

[33] In order for an efficient result to be sustainable, the probability of detection F^{ES} and/or G^{SS} need not to vanish in the limit. We cannot say much about the limit value of these probabilities, since they depend on the exact form of $b^*(\Delta)$ and $\underline{b}^*(\Delta)$ which are unknown to us. These simple conditions on the informativeness of the public signals are sufficient to keep players with incentives to provide effort in the limit. Fudenberg and Levine (2007) discuss the necessity of similar conditions for the existence of an efficient limit equilibrium.

[34] In the context of the partnership game, by measurable distance, we are referring to $2\pi' = |2\pi' - 4\pi'| = |0 - 2\pi'|$.

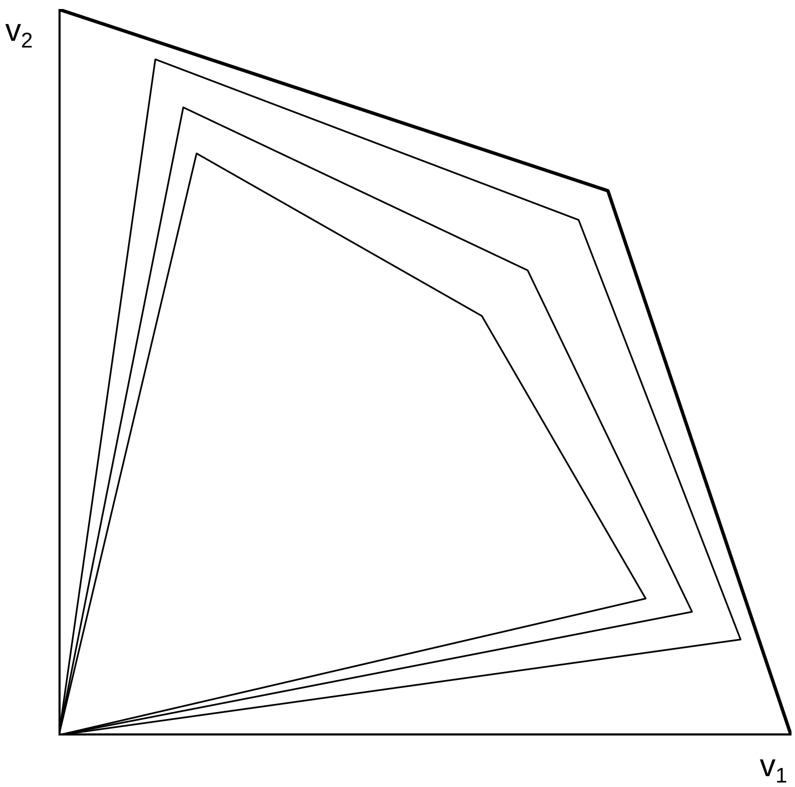

Figure 3: $M^{set}(\Delta, r)$, for $\Delta = 3$ (smallest set) to $\Delta \to 0$ (largest set).

crucial role and makes destruction of value along the equilibrium path converging to zero with Δ. Another important effect that occurs for small Δ is the decrease in the expected immediate gains from a deviation. The reader is referred to the discussion in the end of Section 1.4 above, where these effects are analyzed in some detail.

Figure 3, illustrates the expansion of $M^{set}(\Delta, r)$, for decreasing Δ.

Wrong inference about players actions vanish in the limit. Relevant uncertainty arises only if players cannot observe the public process during some measurable time interval. Then the accumulated sum of infinitesimal normal events may be misleading, which is more likely, the larger the time interval during which the process was left unattended.

The mechanics of the repeated games plays a role in the result. That is, actions are decided by players at the beginning of each period, at the end of the period the state of the process is observed, uncertainty is reset,[35] new actions are taken for the following period

[35] Uncertainty reset is equivalent to saying that either the value of the process is reset or players *actively*

and so on. Under Brownian uncertainty, when we shrink the time interval of this cyclical process to the limit, what Fudenberg and Levine (2007) call "fast play", signals become "almost" perfectly informative about the true profile of actions. Almost, in the sense that we still have some infinitesimal variation, and because we are also not able to determine the identity of the deviator.[36]

The results presented in this paper also hold if we consider other well known Gaussian processes of infinitesimal variation, as for example, the geometric Brownian motion or the Ornstein-Uhlenbeck process. Finally, as discussed in section 1.3, without any extra conditions on the information structure, these results can be generalized to others games with more complex structures.

1.6 Possible Extensions - Some Comments

In this section we briefly discuss two important extensions associated with the methodology presented in this paper. The first is how the results from a continuous time analogue partnership game would differ from the discrete time version. Another important extension would be to see how the approach of the present paper would perform if players have a continuum of actions available. This is important since the methodology of the present paper exploits the jumping effect caused by a discrete deviation.

1.6.1 The Continuous Time

A number of technical issues arise when trying to define a continuous time version of the exercise presented in this paper, namely the limit of a sequence of discrete time games. These problems are related with the time associated with the response to a deviation. For example, if we assume that punishments occur immediately after a deviation at a given time t, in continuous time such moment in time is not well defined (neither the immediate time before). For any $s > t$ we can always find an s' that is smaller than s. One possibility

coordinate. See the discussion on *active coordination* in Section 1.3.

[36] Information about the identity of the deviator is on the basis of strong folk theorems under perfect monitoring, see for example Fudenberg and Maskin (1986), as well as Fudenberg, Levine and Maskin (1994) for a folk theorem under imperfect public monitoring.

to deal with this situation is to assume that punishments occur on the same time as the deviation. In technical terms, such assumption can solve the problem of the denseness of the real numbers. In game theoretical terms, such is not compatible with the assumption of nonanticipatory strategies. An information structure with these characteristics does not correspond to a well-posed problem with strategic interaction. Some of these issues, and others, have been addressed in Simon and Stinchcombe (1989) and Bergin and MacLeod (1993). One possibility discussed in these contributions, is to allow for a ε delay on the punishments. Equilibrium payoffs are found as in a ε-subgame perfect equilibrium (see Fudenberg and Levine (1986)).

Without specific assumptions of this kind we cannot define a continuous time analogue of the problem discussed in this paper. Even though, there is no guarantee that equilibria and strategies in continuous time match with the discrete time version, in particular for more complex equilibrium paths, other than the SSE path.

1.6.2 A Game with a Continuous Action Space

When the action space is discrete in the limit, deviations from the equilibrium path are similar to jumps in the process. Since Brownian paths are continuous but not smooth, such defective behaviour is detected almost surely, in the limit. In this section we briefly discuss the infinitely repeated Cournot game. This game is of interest since it has a continuum of actions available for each player and so deviations can be of infinitesimal magnitude. We will see that the same results presented in the previous Sections keep holding.

We apply the bang-bang result for the strongly symmetric equilibrium in the spirit of Abreu, Pearce and Stacchetti (1986, 1990).[37]

In brief, the stage game expected payoffs (ex-ante) are given by $\pi_i(q_1, q_2) = q_i P(Q)$, where $P(Q)$ is the inverse demand function and $Q = q_1 + q_2$ is the aggregate supply. Firms, always have the possibility of staying out of the market and producing nothing. For simplicity and without loss generality, production costs are zero and firms face no capacity constraints, i.e. $q_i \in [0, \infty)$.

[37] I thank Andrzej Skrzypacz for providing me with the material needed to compute and understand the mechanics of the best SSE payoffs in the Cournot game.

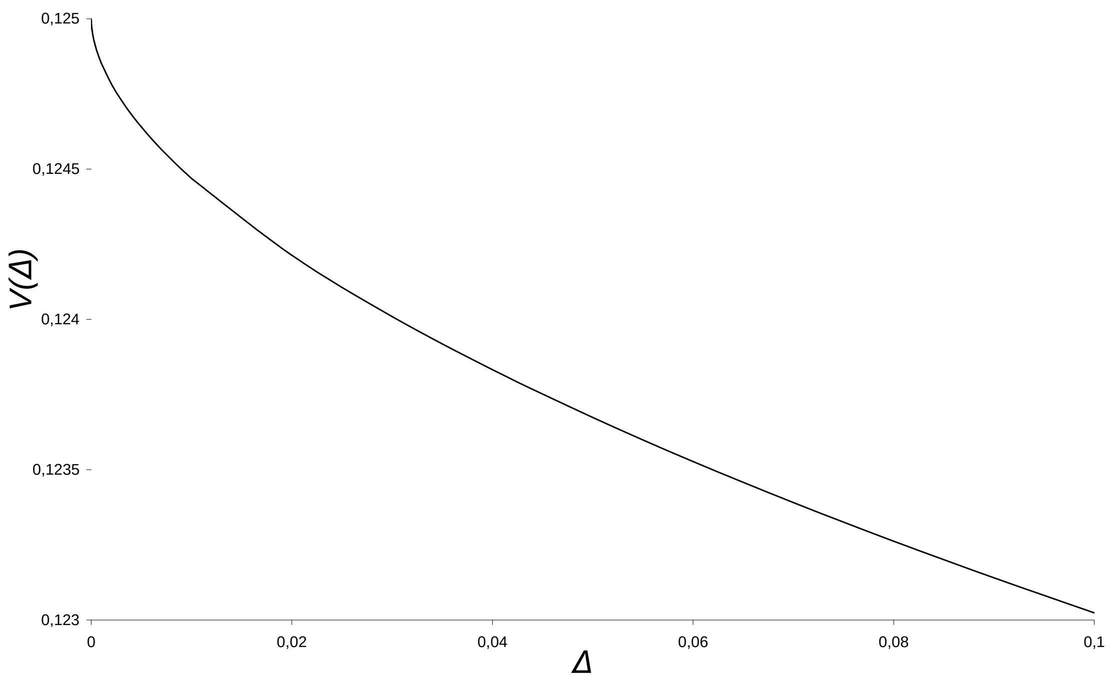

Figure 4: The best SSE in the Cournot game.

The stage game best strongly symmetric equilibrium is achieved when each firm supplies half of the monopoly quantity. Denote this quantity as q_i^M.

The two firms decide their supply quantities simultaneously and independently at moments in time $t = 0, \Delta, 2\Delta...$, and observe the market price (the public signal) at times $t = \Delta, 2\Delta....$ The public signal is given by (1).[38] Now, $y_{t+\Delta}$ is the end of period observed market price, and $y_t \equiv P(Q_t)$ is the initial condition of the process, it reflects the aggregate of individual private supply decisions chosen by each firm at the beginning of the period t.

For the case where $P(Q) = 1 - q_1 - q_2$ and $\sigma = r = 0.1$, Figure 4 shows the value of the best SSE for varying Δ. In particular when Δ becomes small, the best SSE payoff converges to the value 1/8, the payoff of the perfect monitoring 50/50 monopoly split. The numerical approximation suggests that the same efficient limit results shown for the finite action space case (Proposition 7) also holds with a continuous of actions. A monotonic improvement in

[38] The observed state of the ABM price process may take negative values. This feature of the model is irrelevant for the issue we wish to study and there is no loss in generality when considering such a process. A geometric Brownian motion process would be more adequate, but generates other technical problems.

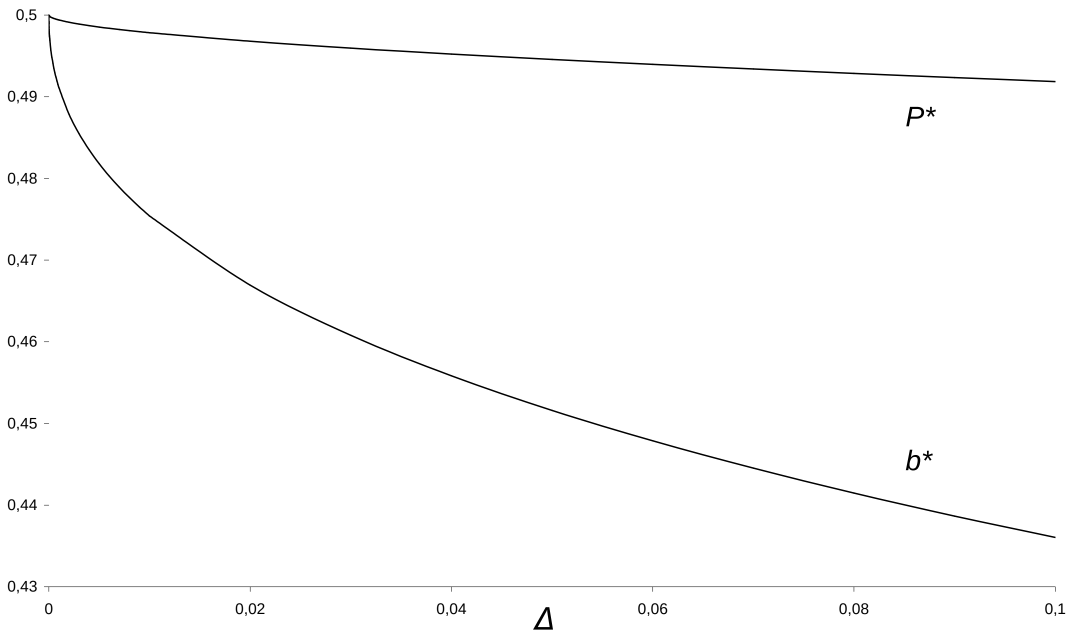

Figure 5: The expected public signal and the optimal threshold value.

the payoffs is clear when Δ becomes small, similar to the statement of Proposition 6.

It is also interesting to contrast the evolution of the optimal threshold b^*, against the expected signal of the process $P(Q^*)$, for varying Δ. Figure 5, shows that as Δ gets small, the threshold value becomes tighter, in a similar fashion to the discrete actions case. In the limit converging to the expected signal associated with the most collusive equilibrium $P\left(Q^M\right) = 1/2$, as can be seen in Figure 5, reflecting the decreasing uncertainty when $\Delta \to 0$. This result is just an extension of Lemma 5, for games with a continuous action space; in the limit the optimal threshold $b^*(\Delta)$ must converge to the expected signal associated with the deviation with less impact on the distribution of the public signals. With a continuous action space, such a deviation is infinitesimal.

The discussion suggests that all the results presented in Section 1.5, also hold when we consider games with a continuous action space. This is the main point that present Section attempts to highlight.

1.7 *Final Comments*

In the simple setting of a repeated partnership game, this paper shows that a folk theorem can be achieved in the limit (the time interval between actions and observation of the process tends to zero), when the public signal observed by the players is the state of an ABM process and the initial conditions associated with different effort profiles are distinct. We also found a monotonic improvement on the payoffs with the monitoring intensity.

Under Brownian uncertainty a degeneracy of the equilibrium in the limit or the opposed efficient scenario as shown here, depends crucially on the modelling adopted. Whereas the former approach tends to fit better in problems where the precision of the signals increases with the time interval between observation of the process, as shown in the work of Fudenberg and Levine (2007, 2009) and Sannikov and Skrzypacz (2007), the present approach is more appropriate in situations where the informativeness of the signals improve with the monitoring intensity.

Reality evolves continuously, however, in economics, and contrary to other sciences, it is hard to think of situations where information is continuously available. An example could be the listed price of very liquid stocks or certain commodities that are available at high frequencies, but not continuously. Nevertheless, if such a possibility were available, this paper shows that the most efficient outcomes would be achieved by continuously monitoring the state of process.

The reason why continuous monitoring of the available information is not seen in real economic problems is because it is costly and may not compensate the potential benefits. In practical applications when a partner continuously monitors the other, she probably cannot devote her time to other activities, such as contributing with effort to the partnership. In practice, what we observe are agents monitoring at discrete moments in time. In many occasions these monitoring events might be random, in the sense that an agent does not know the exact moment in time when the monitor is going to observe the public signal. Osório-Costa (2008) studies repeated game problems of this kind.

Quoting Alchian and Demsetz (1972, p. 780), *"If detecting such behavior were costless,* *neither party would have an incentive to shirk, because neither could impose the cost of his shirking on the other."*

Even though the economic reality apparently finds no place for continuous monitoring, it is important to stress the importance of the results presented here. Although its focus was in the limit, this paper connects, monitoring frequency with their associated payoffs and decision rules. While the impatience level of the players has typically been presented as an exogenous element in the repeated games theory, monitoring frequency has an enormous appeal to be endogenously determined by the problem at hand. This allows repeated games theory in discrete time to study problems in a richer fashion.

1.8 References

1. Abreu, D., P. Milgrom and D. Pearce (1991). "Information and Timing in Repeated Partnerships." Econometrica, 59, 1713-1733.

2. Abreu, D., D. Pearce and E. Stacchetti (1986). "Optimal Cartel Equilibria with Imperfect Monitoring," Journal of Economic Theory, 39, 251-269.

3. Abreu, D., D. Pearce and E. Stacchetti (1990). "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." Econometrica, 58, 1041-1063.

4. Alchian, A. and H. Demsetz (1972). "Production, Information Costs, and Economic Organization." American Economic Review, 62, 777-795.

5. Bergin, J. and W. B. MacLeod (1993). "Continuous Time Repeated Games." International Economic Review, 34, 21-37.

6. Dickinson, D. and M.-C. Villeval (2008). "Does monitoring decrease work effort? The complementarity between agency and crowding-out theories," Games and Economic Behavior, 63, 56–76.

7. Faingold, E., Y. Sannikov (2007). "Reputation Effects and Equilibrium Degeneracy in Continuous-Time Games," mimeo.

8. Frey, B., (1993). "Does monitoring increase work effort? The rivalry between trust and loyalty," Economic Inquiry 31, 663–670.

9. Fudenberg, D. and D. Levine (1986). "Limit Games and Limit Equilibria." Journal of Economic Theory, 38, 261-279.

10. Fudenberg, D. and D. Levine (1994). “Efficiency and Observability with Long-Run and Short-Run Players.” Journal of Economic Theory, 62, 103-135.

11. Fudenberg, D. and D. Levine (2007) “Continuous Time Models of Repeated Games with Imperfect Public Monitoring.” Review of Economic Dynamics, 10(2), 173-192.

12. Fudenberg, D. and D. Levine (2009) “Repeated Games with Frequent Signals.” Quarterly Journal of Economics, 124, 233-265.

13. Fudenberg, D., D. Levine and E. Maskin (1994). “The Folk Theorem with Imperfect Public Information.” Econometrica, 62, 997-1040.

14. Fudenberg, D. and E. Maskin (1986). "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," Econometrica, 54, 533-554.

15. Fudenberg, D. and J. Tirole (1991) Game Theory, MIT Press, Cambridge, MA.

16. Fudenberg, D. and W. Olszewski (2008). "Repeated Games with Asynchronous Monitoring of an Imperfect Signal," mimeo.

17. Green, E. and R. Porter (1984). “Noncooperative Collusion under Imperfect Price Information.” Econometrica, 52, 87-100.

18. Kandori, M. (1992) “The Use of Information in Repeated Games with Imperfect Monitoring.” Review of Economic Studies, 59, 581–593.

19. Mailath, G. and L. Samuelson (2006) Repeated Games and Reputations: Long-run Relationships. Oxford University Press, New York.

20. Osório-Costa, A. M. (2008), "Repeated Games at Random Moments in Time," mimeo.

21. Porter, R. (1983). "Optimal Cartel Trigger Price Strategies." Journal of Economic Theory, 29, 313-338.

22. Prakasa Rao, B.L.S. (1999) Statistical Inference for Diffusion Type Processes. Arnold Publishers and Oxford University Press.

23. Radner, R., R. Myerson, and E. Maskin (1996) "An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria," Review of Economic Studies, 53, 59-69.

24. Sandberg, I. W. (1981) "Global Implicit Function Theorems," IEEE Transactions on Circuits and Systems, 28, 145–149.

25. Sannikov, Y. (2007). "Games with Imperfectly Observable Actions in Continuous Time," Econometrica, 75, 1285–1329.

26. Sannikov, Y. and A. Skrzypacz (2007) "Impossibility of Collusion under Imperfect Monitoring with Flexible Production," American Economic Review, 97, 1794-1823.

27. Sannikov, Y. and A. Skrzypacz (2009) "The Role of Information in Repeated Games with Frequent Actions," Forthcoming in Econometrica.

28. Simon, L. and M. Stinchcombe (1989). "Extensive Form Games in Continuous Time: Pure Strategies," Econometrica 57, 1171-1214.

1.9 Appendix - Proofs of the Lemmas and Propositions.

Proof of Lemma 1. After have solved the system composed by (2), (3) and (4), we obtained $\overline{v}$, $\underline{v}$ and $P = p$, given respectively by (8), (9) and (7).

We want to show that $\overline{v}$ increases monotonically with a decrease in b. Differentiate $\overline{v}$ with respect b we obtain

$$\frac{\partial \overline{v}}{\partial b} = -\left(\pi' - \pi\right) \frac{F^{ES}F_b^{EE} - F^{EE}F_b^{ES}}{\left(F^{ES} - F^{EE}\right)^2},$$

which is always negative, since for the Gaussian distribution, $F^{ES}F_b^{EE} - F^{EE}F_b^{ES} > 0$, for all b. Where, the partial derivatives of F^{EE} and F^{ES} with respect to b are denoted respectively as $F_b^{EE} \equiv \partial F^{EE}/\partial b$ and $F_b^{ES} \equiv \partial F^{ES}/\partial b$. Consequently, we want b smaller as possible but constrained to satisfy (5) and (7). For that we need to know how b changes with p, in particular we are interested in find under which conditions $\partial b/\partial p \leq 0$. By Implicit differentiating (7), we obtain

$$\frac{\partial b}{\partial p} = -\frac{\pi' - \pi}{F_b^{ES}\pi - F_b^{EE}\pi'}.$$

This derivative is negative if

$$F_b^{EE}/F_b^{ES} = e^{\frac{4b\pi' - 12\pi'^2}{2\Delta\sigma^2}} < \pi/\pi'. \tag{12}$$

The RHS is a number between 0 and 1. The LHS is continuous and monotonic in b, when $b = 3\pi'$, it reaches the value 1, which does not satisfy the inequality. When $b \to -\infty$, the LHS goes to 0, which satisfy the inequality. By continuity of the LHS in b, there exist be a value $b \leq 3\pi'$ below which the inequality holds. This value is given by

$$b^p = 3\pi' + \sigma^2\Delta \ln\left(\pi/\pi'\right)/2\pi',$$

and it is an asymptote line of $\partial b/\partial p$.

The value b^p is also the unique value that minimizes the LHS of (7) with respect to b. To show it, notice that (12) when holding with equality is equivalent to $\partial P/\partial b = 0$. The second order condition is

$$\frac{\partial^2 P}{\partial b^2} = -\frac{F_{bb}^{ES}\pi - F_{bb}^{EE}\pi'}{\pi' - \pi}, \tag{13}$$

where, the second derivatives of F^{EE} and F^{ES} with respect to b are denoted respectively as $F_{bb}^{EE} \equiv \partial^2 F^{EE}/\partial b^2$ and $F_{bb}^{ES} \equiv \partial^2 F^{ES}/\partial b^2$. This derivative is always strictly positive. To see it, notice that $F_{bb}^{EE}(b,\Delta) = -(b - 4\pi') F_b^{EE}/\Delta\sigma^2$ and $F_{bb}^{ES}(b,\Delta) = -(b - 2\pi') F_b^{ES}/\Delta\sigma^2$. Replace these expressions in (13) and use the first order condition given by (12) when holding with equality, to obtain

$$\frac{\partial^2 P}{\partial b^2} = \frac{2\pi'^2 F_b^{EE}}{\Delta\sigma^2(\pi' - \pi)} > 0,$$

which is always strictly positive. So, b^p is the value that minimizes P, but notice that this is an unconstrained minimum. For that reason, P might take a negative value at b^p, or even be larger than 1. In either case, condition (7) is not satisfied since $p \in [0,1]$. While in the latter case we cannot enforce the profile $(1,1)$, in the former, we can decrease b below b^p until P be at least equal to 0, i.e. feasible, because $\partial b/\partial p < 0$ for $b < b^p$. But in this direction, the value of b that maximizes $\overline{v}$ and satisfy (7) can be pushed even lower, increase P until 1. The value $p = 1$ is the optimal feasible choice, that keeps (7) holding with equality. Such value of $b \leq b^p$ is the optimal threshold b^* and the optimal punishment is $\underline{v} = 0$, i.e. perpetual punishment. ■

Proof of the Lemma 3. When $\Delta \in \left(0, \overline{\Delta}\right)$ we always have $b^* < b^v < b^p$, with $b^p \equiv \arg\min_b P$, $b^v \equiv \arg\max_b \underline{v}$ and $b^* \equiv \{\arg\max_b \overline{v} : \underline{v} = 0\}$. Then, at $\Delta = \overline{\Delta}$, $P(b^p) = 1 \Longleftrightarrow \underline{v}(b^p) = 0$ and $P(b^v) = 1 \Longleftrightarrow \underline{v}(b^v) = 0$, implying that the three optimization problems have to satisfy the same constraint. Consequently, $b^p\left(\overline{\Delta}\right) \equiv \arg\min_b P = b^v\left(\overline{\Delta}\right) \equiv \arg\max_b \underline{v}$.[39]

Since P is strictly convex, the $\min_b P$ gives a first order condition, that can be written as

$$F_b^{EE}/F_b^{ES} = \pi/\pi',$$

where b^p given in (11) solve the equality. At $\Delta = \overline{\Delta}$, P has a minimum with respect to b at some point, then $\underline{v}$ must have a maximum with respect to b at the same point. It can be shown that $\underline{v}$ is strictly concave with a unique maximum, to simplify lets assume it. After some arrangements the first order condition of $\max_b \underline{v}$ can be written as

$$\delta\left[F^{ES}F_b^{EE} - F^{EE}F_b^{ES}\right] = (1-\delta)\left[F_b^{ES} - F_b^{EE}\right],$$

where b^v solves the equality. Because at $\overline{\Delta}$, we must have $b^p = b^v$ the two first order conditions must be satisfied at the same value. Manipulating both equalities we obtain

$$F^{ES}(b^p)\,\pi - F^{EE}(b^p)\,\pi' = (1-\delta)\left(\pi' - \pi\right)/\delta,$$

which is (10), with b replaced by b^p, given in (11). The value of Δ that solves this equality is $\overline{\Delta}$. ■

[39] Otherwise, for $\Delta > \overline{\Delta}$, we have $b^* > b^v > b^p$. However, in this region neither of these value is defined in any feasible sense, since (10) cannot be satisfied anymore.

Proof of Lemma 4. We extend the usual local implicit function theorem to hold in the convex interval $\left(0, \overline{\Delta}\right)$. Lemma 3 tell us how to find the value $\overline{\Delta}$. Moreover, for bounded σ and r, a value of $\overline{\Delta} > 0$ always exist. Rewrite the equality (10) in the following way and denote it as $I(b, \Delta)$, i.e.

$$I(b, \Delta) \equiv F^{ES}\pi - F^{EE}\pi' - (1-\delta)\left(\pi' - \pi\right)/\delta. \tag{14}$$

Since F^{EE}, F^{ES} and δ are continuous and differentiable with respect to $\Delta \in \left(0, \overline{\Delta}\right)$ and $b \in (-\infty, b^p]$, so thus the mapping $I(b, \Delta)$ is continuous and differentiable.

From, Lemma 1 and Lemma 3, for any $\Delta_0 \in \left(0, \overline{\Delta}\right)$ there is exactly one $b_0 \in (-\infty, b^p]$ such that $I(b_0, \Delta_0) = 0$. If $\Delta_0 > \overline{\Delta}$ there is no such value of b_0, and if $b \in \mathbb{R}$, b_0 is not unique, because P given by (7), is strictly convex with a minimum at b^p. Consequently, point (i) of Sandberg's (1981, p. 146) global implicit function theorem is satisfied.

Additionally, we need to verify that $I(b, \Delta)$ is locally solvable in the neighborhood of the point (b_0, Δ_0) and then by continuity of $b(\Delta)$, Sandberg's theorem holds for all $\Delta \in \left(0, \overline{\Delta}\right)$. The condition for local solvability is $\partial I(b_0, \Delta_0)/\partial b \neq 0$, implying that $I(b_0, \Delta_0) = 0$. The differentiability condition is written as

$$F_b^{ES}\pi - F_b^{EE}\pi' \neq 0,$$

where, the partial derivatives of F^{EE} and F^{ES} with respect to b are denoted respectively as $F_b^{EE} \equiv \partial F^{EE}/\partial b$ and $F_b^{ES} \equiv \partial F^{ES}/\partial b$. In the proof of Lemma 1 we have seen that this equation is satisfied for all $b_0 \neq b^p$, with b^p given by (11). However, to improve on the value of the best SSE payoff we are interested in $b_0 \leq b^p$, implying that $\partial I(b_0, \Delta_0)/\partial b \geq 0$. Notice, that $b_0 = b^p$ (and consequently $\partial I(b_0, \Delta_0)/\partial b = 0$) only at $\Delta = \overline{\Delta}$. But for all $\Delta_0 \in \left(0, \overline{\Delta}\right)$, we always have, $b < b^p$ (and consequently $\partial I(b_0, \Delta_0)/\partial b > 0$). Then for each point inside the interval $\left(0, \overline{\Delta}\right)$ there is a unique continuous differentiable function $b(\Delta)$ on an open ball around (b_0, Δ_0) that locally satisfies $I(b, \Delta) = 0$.

Now, apply Sandberg's implicit function theorem; by continuity of $b(\Delta)$, for each $S \in A$ there is a $T \in B$, where A and B are families of compact subsets of $\left(0, \overline{\Delta}\right)$ and $(-\infty, b^p)$ respectively, $b(S)$ is compact as well and belongs to T. Then, there is a unique, continuous and differentiable function $b^*(\Delta)$ such that $I(b^*, \Delta) = 0$ for all $\Delta \in \left(0, \overline{\Delta}\right)$. ■

Proof of Lemma 5. Since $\partial\overline{v}/\partial b < 0$ is always true, lower the value b, larger the value $\overline{v}$, see the proof of Lemma 1. Moreover, by Proposition 6, $\overline{v}^*$ improves monotonically as Δ gets small, the largest value $\overline{v}^*$ is reached in the limit $\Delta \to 0$. Our problem is then to find the lowest feasible value of b that maximizes $\overline{v}$, in the limit, and on same time satisfies (10) and its derivatives with respect to Δ. We rewrite (10) here

$$F^{ES}\pi - F^{EE}\pi' = (1-\delta)\left(\pi' - \pi\right)/\delta.$$

The RHS goes to 0 with Δ. So in the limit the LHS must also go to 0 with Δ. The upper bound $b^p \to 3\pi'$ when $\Delta \to 0$, so we don't need to consider values of $b > 3\pi'$. For $b \to x \in (2\pi', 3\pi']$, we have $F^{ES} \to 1$ and $F^{EE} \to 0$. Then LHS goes to $\left(\pi - \pi^N\right) > 0$, the enforceability condition (10) is satisfied but with slack. So the limit value of $b^*(\Delta)$ must be lower.

For $b \to x \in (-\infty, 2\pi')$, we have $F^{ES} \to 0$ and $F^{EE} \to 0$, both the LHS and the RHS goes to 0, but we need to check for indetermination, since the LHS might go faster to 0 than the RHS. Differentiate both sides with respect to Δ we obtain

$$-\frac{b-2\pi'}{2\Delta^{3/2}\sigma\sqrt{2\pi}}e^{-\frac{(b-2\pi')^2}{2\Delta\sigma^2}}\pi + \frac{b-4\pi'}{2\Delta^{3/2}\sigma\sqrt{2\pi}}e^{-\frac{(b-4\pi')^2}{2\Delta\sigma^2}}\pi' = r\left(\pi' - \pi\right)/\delta. \tag{15}$$

When $b \to x \in (-\infty, 2\pi')$, the limit on LHS is smaller than the limit on the RHS, i.e. $0 < r(\pi' - \pi)$. Meaning that the LHS of (10) is smaller than the RHS, i.e. we lose enforceability. So, when $\Delta \to 0$ we cannot have $b \to x \in (-\infty, 2\pi')$.

Consequently, we must have $\lim_{\Delta\to 0} b^*(\Delta) = 2\pi'$. The order at which it converge must be such that (10) and (15) hold with equality. ■

Proof of Proposition 6. Start by defining $\overline{v}^*(\Delta) = \{\max_{b\leq b^p}\overline{v}(b,\Delta) : I(b,\Delta) = 0\}$ and write the Lagrangian $\mathcal{L}(b,\Delta) = \overline{v}(b,\Delta) - \lambda I(b,\Delta)$. By Lemma (4) the solution $b^*(\Delta)$ is a continuous and differentiable function of Δ; assume the same holds for the Lagrangian multiplier λ. Then by the envelope theorem for constrained maximization problems we can write $\partial\overline{v}^*(\Delta)/\partial\Delta = \partial\mathcal{L}(b^*,\Delta)/\partial\Delta$. Our goal is then to show that

$$\frac{\partial\overline{v}^*(\Delta)}{\partial\Delta} = \frac{\partial\overline{v}(b^*,\Delta)}{\partial\Delta} - \lambda\frac{\partial I(b^*,\Delta)}{\partial\Delta} < 0 \tag{16}$$

where λ is obtained from solving $\partial\mathcal{L}(b,\Delta)/\partial b = 0$. Expression (16) has the following three components, which we develop to

$$\begin{aligned}
\frac{\partial\overline{v}(b^*,\Delta)}{\partial\Delta} &= -\frac{F^{ES}F^{EE}_{\Delta} - F^{EE}F^{ES}_{\Delta}}{\left(F^{ES}-F^{EE}\right)^2}\left(\pi'-\pi\right), \\
\frac{\partial I(b^*,\Delta)}{\partial\Delta} &= F^{ES}_{\Delta}\pi - F^{EE}_{\Delta}\pi' - \frac{r}{\delta}\left(\pi'-\pi\right), \\
\text{and }\ \lambda &= -\frac{\left(F^{ES}F^{EE}_{b} - F^{EE}F^{ES}_{b}\right)\left(\pi'-\pi\right)/\left(F^{ES}-F^{EE}\right)^2}{F^{ES}_{b}\pi - F^{EE}_{b}\pi'}.
\end{aligned}$$

Where, we have denoted $F^{EE}_{\Delta} \equiv \partial F^{EE}/\partial\Delta$, $F^{ES}_{\Delta} \equiv \partial F^{ES}/\partial\Delta$, $F_b \equiv \partial F^{EE}/\partial b$ and $F^{ES}_{b} \equiv \partial F^{ES}/\partial b$ when evaluated at $b = b^*$. Replacing these expressions in (16), and after some algebra we obtain

$$\left(F^{ES}\pi - F^{EE}\pi'\right)\left(F^{ES}_{b}F^{EE}_{\Delta} - F^{EE}_{b}F^{ES}_{\Delta}\right) > -\frac{r}{\delta}\left(\pi'-\pi\right)\left(F^{ES}F^{EE}_{b} - F^{EE}F^{ES}_{b}\right),$$

using the expression (10), we can simplify further to get

$$(1-\delta)\left(F^{ES}_{b}F^{EE}_{\Delta} - F^{EE}_{b}F^{ES}_{\Delta}\right) > -r\left(F^{ES}F^{EE}_{b} - F^{EE}F^{ES}_{b}\right).$$

Notice that $F^{EE}_{\Delta} = -\left(b^* - 4\pi'\right)F^{EE}_{b}/2\Delta$ and $F^{ES}_{\Delta} = -\left(b^* - 2\pi'\right)F^{ES}_{b}/2\Delta$, then we are left with

$$(1-\delta)F^{EE}_{b}F^{ES}_{b}\pi'/\Delta > -r\left(F^{ES}F^{EE}_{b} - F^{EE}F^{ES}_{b}\right).$$

The LHS is always positive. The RHS is always negative, since for the Gaussian distribution $F^{ES}F^{EE}_{b} - F^{EE}F^{ES}_{b} > 0$ for any $\Delta > 0$. ■

Proof of Proposition 7. Let $x : Y \to \mathbb{R}^2$ be the vector of normalized continuations, with generic element $x_i(y)$ defined as

$$x_i(y) \equiv \frac{\delta}{1-\delta}\left(w_i(y) - v_i\right), \tag{17}$$

with $y \in Y$ and where $w_i : Y \to \mathbb{R}$ is the usual player's i continuation payoff. With just two signals $Y \equiv \left\{\underline{y}, \overline{y}\right\}$, denote $x_i\left(\underline{y}\right) \equiv \underline{x}_i$ and $x_i(\overline{y}) \equiv \overline{x}_i$ respectively.

Our goal is to maximize the score $k^*(a,\lambda) = \max_{x(y)} \lambda.v$ by choosing a vector $x(y)$ that enforces a in the direction $\lambda \equiv (\lambda_1, \lambda_2)$, i.e. satisfying

$$v_i = \pi_i(a) + E\left[x_i(y)\,|a\right],\ \forall i \tag{18}$$

$$v_i \geq \pi_i\left(a'_i, a_{-i}\right) + E\left[x_i(y)\,|\left(a'_i, a_{-i}\right)\right],\ \forall a'_i \in A_i,\ \forall i \tag{19}$$

and

$$0 \geq \lambda.x\left(y\right), \ \forall y \in Y. \tag{20}$$

The transformation (17) allow us to enforce action profiles independently of the discount factor δ. However, we have no guarantee that the resulting continuation values are either in the set of PPE $E^{set}\left(\Delta, r\right)$, or even feasible (In the limit $M^{set}\left(\Delta\right) = \lim_{r\to 0} E^{set}\left(r, \Delta\right)$). Let $H\left(\lambda, k^*\right) \equiv \left\{v \in \mathbb{R}^2 : \lambda.v \leq k^*\left(a, \lambda\right)\right\}$ be the half-space that decomposes the payoff v in the direction λ with the maximum score $k^*\left(a, \lambda\right)$.

Suppose $\lambda_i > 0$, for $i = 1, 2$. We start by enforcing the profile $a = (1, 1)$. By choosing $\underline{x}_i = -\left(\pi' - \pi\right) / \left(F^{ES} - F^{EE}\right)$ and $\overline{x}_i = 0$, we make (19) hold with equality, but not (20).[40] We obtain

$$\lambda.v \leq k^*\left(\left(1,1\right), \lambda\right) = \left(\lambda_1 + \lambda_2\right)\left(\pi - \frac{F^{EE}}{F^{ES} - F^{EE}}\left(\pi' - \pi\right)\right) = \left(\lambda_1 + \lambda_2\right)\overline{v}.$$

To enforce the profile $a = (1, 0)$, we set $\overline{x}_1 = \left(\pi' - \pi\right)/2\left(G^{SS} - G^{ES}\right)$, $\underline{x}_1 = -\left(\pi' - \pi\right)/2\left(G^{SS} - G^{ES}\right)$ and $x_2\left(y\right) = -\lambda_1 x_1\left(y\right)/\lambda_2$. Both (19) and (20) hold with equality, and we have

$$\lambda.v = k^*\left(\left(1,0\right), \lambda\right) = -\lambda_1\left(\pi' - \pi\right) + \lambda_2 \pi'.$$

Now, suppose $\lambda_2 > 0 > \lambda_1$. (i) Consider first $\lambda_2 \geq -\lambda_1 > 0$. The profile $a = (1, 0)$ can be enforced with both (19) and (20) holding with equality. Setting $\overline{x}_1 = G^{ES}/\left(G^{SS} - G^{ES}\right)$, $\underline{x}_1 = -\left(1 - G^{ES}\right)/\left(G^{SS} - G^{ES}\right)$ and $x_2\left(y\right) = -\lambda_1 x_1\left(y\right)/\lambda_2$ for $y = \underline{y}, \overline{y}$.

(ii) Consider now the case $-\lambda_1 > \lambda_2 > 0$. Constraint (19) holds with equality if we choose $\underline{x}_1 = \underline{x}_2 = 0$, $\overline{x}_1 = \left(\pi' - \pi\right)/\left(G^{SS} - G^{ES}\right)$ and $\overline{x}_2 = \left(\pi' - \pi\right)/\left(F^{ES} - F^{EE}\right)$, then

$$\lambda.v \leq k^*\left(\left(1,0\right), \lambda\right) = -\lambda_1\left(\pi' - \pi\right) + \lambda_2\pi' + \lambda_1\left(1 - G^{ES}\right)\frac{\pi' - \pi}{G^{SS} - G^{ES}} + \lambda_2\left(1 - F^{ES}\right)\frac{\pi' - \pi}{F^{ES} - F^{EE}}. \tag{21}$$

Since $\left(1 - G^{SS}\right) > 0$ for large $-\lambda_1 > 0$ we obtain $k^*\left(\left(1,0\right), \lambda\right) < k^*\left(\left(0,0\right), \lambda\right) = 0$. By symmetry we can get the other boundaries of the M^{set} set.

The condition that establishes that the best strongly symmetric equilibrium $\overline{v}$ is larger than the best symmetric equilibrium is $\overline{v} > \pi/2$. Developing further we obtain

$$F^{ES}\pi - F^{EE}\pi' > \pi' - \pi. \tag{22}$$

[40] If (20) holds with equality, we say that $x\left(y\right)$ for $y = \underline{y}, \overline{y}$, orthogonally enforces a in the direction λ. In this case a transfer of value between the players at prices λ is possible.

Otherwise, the symmetric equilibrium obtained by alternating play, between $(1,0)$ and $(0,1)$ would be better.

So far, we have found the maximum score $k^*(a,\lambda)$, in a direction λ that enforces the profile a. Now, we can write the half-space $H(\lambda,k^*) \equiv \left\{v \in \mathbb{R}^2 : \lambda.v \leq k^*(a,\lambda)\right\}$, that decomposes the payoff v in the direction λ, with associated maximum score $k^*(a,\lambda)$. Given, the graph of the set of feasible and individual rational payoffs and by symmetry, we have two relevant half-spaces.

Suppose that (22) holds, then the upper left boundary of the set $M^{set}(\Delta,r)$ is given by the line (21) passing through the payoff point $(0,0)$, i.e. $v_2 \leq -\lambda_1 v_1/\lambda_2$ satisfying

$$-\lambda_1\left(\pi'-\pi\right)+\lambda_2\pi'+\lambda_1\left(1-G^{ES}\right)\frac{\pi'-\pi}{G^{SS}-G^{ES}}+\lambda_2\left(1-F^{ES}\right)\frac{\pi'-\pi}{F^{ES}-F^{EE}}=0.$$

Which we can solve to obtain the condition on the direction λ, i.e.

$$-\frac{\lambda_1}{\lambda_2}=\frac{\left(G^{SS}-G^{ES}\right)\left(\left(1-F^{EE}\right)\pi'-\left(1-F^{ES}\right)\pi\right)}{\left(\pi'-\pi\right)\left(1-G^{SS}\right)\left(F^{ES}-F^{EE}\right)}.$$

The upper left boundary is then given by

$$v_2 \leq \frac{\left(G^{SS}-G^{ES}\right)\left(\left(1-F^{EE}\right)\pi'-\left(1-F^{ES}\right)\pi\right)}{\left(\pi'-\pi\right)\left(1-G^{SS}\right)\left(F^{ES}-F^{EE}\right)}v_1. \tag{23}$$

The upper right boundary is given by the line that pass through the points $\left(-\left(\pi'-\pi\right),\pi'\right)$ and $(\overline{v},\overline{v})$. Consequently, the direction λ must satisfy

$$\left(\lambda_1+\lambda_2\right)\overline{v}=-\lambda_1\left(\pi'-\pi\right)+\lambda_2\pi',$$

i.e.

$$-\frac{\lambda_1}{\lambda_2}=-\frac{\pi'-\overline{v}}{\overline{v}+\pi'-\pi}.$$

Then, the upper right boundary on the limit set is written as

$$v_2 \leq \frac{2\pi'-\pi}{\overline{v}+\pi'-\pi}\overline{v}-\frac{\pi'-\overline{v}}{\overline{v}+\pi'-\pi}v_1. \tag{24}$$

The interception of the lines (23) and (24) occurs at the point

$$\left(v_1^u,v_2^u\right)=\left(\frac{\frac{2\pi'-\pi}{\overline{v}+\pi'-\pi}\overline{v}}{\frac{\pi'-\overline{v}}{\overline{v}+\pi'-\pi}+\frac{\left(G^{SS}-G^{ES}\right)\left(\left(1-F^{EE}\right)\pi'-\left(1-F^{ES}\right)\pi\right)}{\left(\pi'-\pi\right)\left(1-G^{SS}\right)\left(F^{ES}-F^{EE}\right)}},\frac{\frac{\left(G^{SS}-G^{ES}\right)\left(\left(1-F^{EE}\right)\pi'-\left(1-F^{ES}\right)\pi\right)}{\left(\pi'-\pi\right)\left(1-G^{SS}\right)\left(F^{ES}-F^{EE}\right)}\frac{2\pi'-\pi}{\overline{v}+\pi'-\pi}\overline{v}}{\frac{\pi'-\overline{v}}{\overline{v}+\pi'-\pi}+\frac{\left(G^{SS}-G^{ES}\right)\left(\left(1-F^{EE}\right)\pi'-\left(1-F^{ES}\right)\pi\right)}{\left(\pi'-\pi\right)\left(1-G^{SS}\right)\left(F^{ES}-F^{EE}\right)}}\right).$$

By symmetry we obtain the lower right point $\left(v_1^l, v_2^l\right) = \left(v_2^u, v_1^u\right)$. Together with $(\overline{v}, \overline{v})$ and $(0,0)$, when (22) holds, these points are enough to characterize the limit set $M^{set}(\Delta, r)$.

We want now to take the limit $\Delta \to 0$ of these points. In the Lemmas 1, 3, 4 and 5, we developed a great knowledge about the threshold that sustain the best SSE payoff. In particular, in the limit it must converge to $2\pi'$ from below, moreover any threshold in the region $(2\pi', 3\pi')$ is also feasible, since enforceability condition keeps holding. About the optimal threshold associated with the an asymmetric path starting with the profile $(1,0)$, it must converge to 0 following the discussion after Lemma 5. To simplify and in order to avoid issues related with the order of convergence, it is enough to chose some $b \in (2\pi', 3\pi')$ and some $\underline{b} \in (0, \pi')$, which are feasible and admissible choices. In the limit, enforceability conditions holds with slack, but we still obtain perfectly informative signals, i.e. when $\Delta \to 0$, we have $F^{EE} \to 0$, $F^{ES} \to 1$, $G^{ES} \to 0$ and $G^{SS} \to 1$. Consequently $\overline{v} \to \pi$, $v_1^u = v_2^l \to 0$ and $v_2^u = v_1^l \to (2\pi' - \pi)\,\pi/\pi'$, which are the critical points of the set of feasible and individual rational payoffs. The choices b and $\underline{b}$ are independent of r, for that reason we don't need to take $r \to 0$ to obtain the folk-theorem, because that effect operates on the discount factor through $\Delta \to 0$. Moreover, condition (22) holds in the limit, since by assumption $2\pi > \pi'$, proving the result. ■

CHAPTER II

REPEATED GAMES AT RANDOM MOMENTS IN TIME

2.1 Introduction

Many economic situations where agents repeatedly interact are studied as repeated games. While the research in repeated games has focused exclusively on the case where the repetition of the stage game is deterministic at equally spaced moments in time, this paper considers the possibility that the stage game is played at random and not at equally spaced moments in time. Many economic interactions of interest are in fact similar to repeated games, but not necessarily repeated with some known frequency. Players do not know and do not control, ex-ante, the moments in time in which they are going to play again.

For example, during a working day, an agent might not know exactly at which moment he is going to be monitored by his superior. Even after being monitored, the agent still holds the same uncertainty, since the superior might visit him again at any instant after the last monitoring event. Similarly, two colluded firms need not adjust their actions to the end of a calendar day/month. Exogenous variations, such as the market price, demanded quantities, variations in input price, entry prospectus of new competitors, or the arrival of other types of information, might require these firms (through explicit communication or by common observation of the same events/signals) to adjust their actions more than once during a given day/month, or even not to do it at all.

Accepting this possibility, from a technical point of view, we ask what changes have to be introduced in the usual repeated games methodology in order to accommodate such a possibility? In particular, from the perspective of efficiency, we are concerned with how the payoffs change when there is uncertainty about the time repetitions of the stage game. This paper attempts to answer these questions, not only when the monitoring is perfect, but also when it is public and imperfect.[1][2]

We show that, under perfect monitoring, uncertainty about the repetitions of the stage game improves payoffs in comparison with the usual deterministic case. This is always true when players' discount function is convex in the time domain. Even though their true discount rate remains unchanged, players' decisions are based on a smaller discount rate, which we call the *expected discount factor effect*.

Surprisingly, when the monitoring is imperfect but public, the result does not generalize in the same way. The positive effect of the players' discounting it is not always sufficient to compensate for the adverse effects on the distribution of the public signals for all frequencies of play. Nonetheless, we establish conditions under which random monitoring allows efficiency gains on the value of the best *Strongly Symmetric Equilibrium* (SSE henceforth), when compared with the deterministic approach.

We end the paper with a set of examples that attempt to illustrate the effects of random monitoring in terms of efficiency.

It is not the goal of this paper to write a general theory of repeated games played at random moments in time, rather to present and discuss a set of issues with particular relevance for applied work.

The study of random monitoring, as done in the present paper, would not be possible without the recent advances in the theory of frequent monitoring. After the seminal work of Abreu, Milgrom and Pearce (1991), renewed interest in frequent monitoring has re-emerged, in particular due to Sannikov (2007).[3] More similar to the former work, Fudenberg

[1] When monitoring is imperfect but public, player commonly observes noisy public signals about other players' actions. For example, Radner, Myerson and Maskin (1986) consider the case where the output of a partnership is a noisy signal of players' actions. Green and Porter (1984) and Porter (1983) look at the case where the market price is an imperfect signal of the quantities supplied by firms. See Fudenberg and Tirole (1991) and Mailath and Samuelson (2006) for complete surveys of the problems and methods used to solve repeated games.

[2] The case where the stage game is repeated at unknown and not equally spaced moments in time, we call *random/stochastic monitoring*. When the stage game is repeated at known and equally spaced moments in time, we call it *deterministic monitoring*. These concepts should not be confused with perfect or imperfect monitoring. Perfect monitoring can be either random or deterministic. The same happens with imperfect monitoring.

[3] In the same spirit, studying games in continuous time see Faingold and Sannikov (2007) and Faingold (2006).

and Levine (2007, 2009), Osório-Costa (2008) and Sannikov and Skrzypacz (2007, 2009), study the limit of a sequence of discrete time games parameterized by Δ (the time interval between repetitions of the stage game). These papers were mainly concerned with the limit case, although they sowed the seeds for studying repeated games with arbitrary monitoring intensities.

In the present paper we do not focus on a particular monitoring intensity, but look for the entire spectrum of monitoring frequencies where repeated play can improve over the stage game static Nash payoffs.

We can divide the limit results into two types, depending on whether we improve on the stage game static Nash or not. In the latter case, Abreu, Milgrom and Pearce (1991) show that the value of the best strongly symmetric equilibrium degenerates at the limit when the realizations of the public process represent bad news. The lack of observed public signals becomes infinitely likely at the limit. Fudenberg and Levine (2007, 2009) and Sannikov and Skrzypacz (2007) present similar limit results when the public signal is Brownian rather than Poisson. Their degeneracy effect is due to a degradation of the information content of the public signals for high monitoring intensities.

Not all results obtained point to a degeneracy. When the realizations of the public process are interpreted as bad news, Abreu, Milgrom and Pearce (1991) have shown that equilibrium payoffs above the static Nash, but not fully efficient, can be sustained in the limit. It is, however, at the limit where the best results can be obtained. Under Brownian signals, Fudenberg and Levine (2007) and Osório-Costa (2008) show that full efficiency can emerge at the limit. In the latter paper, players control the drift of the process and different action profiles have different initial conditions associated. In the former contribution, a deviation increases the volatility of the process.[4] In either case, in the limit, signals become perfectly informative about players' actions. Asymptotically, public monitoring converges to perfect monitoring.[5]

[4] Fudenberg and Levine (2007), also show that if a deviation has the inverse effect on the noise parameter it is possible to obtain payoffs above the static Nash, but not fully efficient.

[5] In Section 2.5, we provide numerical illustrations of these results.

Fudenberg and Olszewski (2009) present an interesting model that shares some similarities with the present paper. They study a repeated game with stochastic asynchronous monitoring which leads to a private monitoring problem. However, they focus on the limit case. They show that at the limit, synchronous and asynchronous monitoring are equivalent if the signals are Poisson. However, when the signals are Brownian, in some cases, the limit value of the asynchronous games might be lower.

In the present paper, monitoring is stochastic and synchronous, and we are interested in more general monitoring intensities as well as in perfect and public monitoring information structures. To the best of my knowledge, this is the first paper that studies the implications of the uncertainty on the repetitions of the stage game in infinitely repeated games.

Stochastic repetitions of the stage game (with a Poisson process) are also addressed in Faingold (2006). Such an assumption is motivated by technical reasons; Faingold is mainly concerned with limit reputational effects in games with imperfect monitoring.

The dynamic uncertainty aspect of the present model is novel, not only in the repeated games literature, but also in contract and agency problems. Nonetheless, we can establish some links with other previous works.

This paper shares similarities with the literature on stochastic monitoring or costly state verification in static settings. Starting with the seminal work of Townsend (1979), where he found that random verification can be Pareto superior to deterministic monitoring in most cases, where the monitoring threat, and not necessarily the monitoring action, are sufficient to provide players with incentives. Border and Sobel (1987) characterize efficient stochastic and self-enforcing monitoring, assuming commitment.[6] A repeated version of Border and Sobel's model is studied by Monnet and Quintin (2005). The difference from the present paper is that the monitoring uncertainty occurs in the stage game, that is played at equally spaced moments in time.

[6] Optimal stochastic monitoring strategies in static environments are also studied by Mookherjee and Png (1989), who look at the case of risk-aversion, and Bernanke and Gertler (1989), who study an investment problem where perfect information can be achieved with costly monitoring activity. See also, Khalil (1997), which shows that the lack of commitment leads to a higher probability of monitoring.

Because monitoring is in general a costly activity, once we identify when random monitoring is superior to deterministic monitoring, or the other way around, we can select the monitoring technology that achieves larger payoffs for the same monitoring frequency (i.e. the same expected costs), or from another perspective, we can fix the payoff and search for the less costly monitoring structure.[7]

The paper is organized as follows. Section 2.2 presents the repeated game model and some general notation. It also extends the notion of equilibrium for random monitoring settings. Special attention is given to the expected discount factor effect. Section 2.3 focuses on perfect monitoring and the first results are presented. Section 2.4 studies the public monitoring case. We look at the distribution of the public signals and we characterize the value of the best SSE payoff. Finally we establish the conditions for efficiency gains under random monitoring. In Section 2.5, we illustrate our findings for some important models in frequent monitoring. Section 2.6 concludes. The proofs of the results presented in the main text can be found in Appendix 2.8.

2.2 The Model and the Expected Discount Factor

In this section, we describe the repeated game model that can harbour both the deterministic and random monitoring cases.

At moments in time $t_0, t_1, t_2, ...$, each player $i \in N \equiv \{1, 2, ..., n\}$ chooses an action a_i from some finite action space A_i. Where $t_0 = 0$ is the known moment in time when players play for the first time, t_1 is the moment in time when players play for the second time, and so on. Denote $A = \times_{i=1}^{n} A_i$ as the set of action profiles endowed with the product topology of the individual action spaces, with generic element $a = (a_1, ..., a_n)$ denoting a profile of actions. Actions are taken at the beginning of each period, but payoffs are collected at the end. It is also at these moments in time $t_1, t_2, ...$, that the resulting profile of actions chosen

[7] Under random monitoring, the monitoring events are *i.i.d.* random variables with support in some interval. When, in expected terms, the time length between observations matches the deterministic and equally spaced case; we use the terminology "same monitoring frequency (or intensity)". In expected terms, the monitoring events happens the same number of times. This is important when monitoring is costly, because the expected costs will be the same.

at the beginning of each period is observed, perfectly or with noise.[8]

When monitoring is deterministic, let $t_k - t_{k-1} \equiv \Delta$, with $k \geq 1$. In the case where monitoring is random, $t_k - t_{k-1} \equiv x$ is a random variable following some distribution $G(x)$. The support of this distribution, is chosen such that $E_X(t_k - t_{k-1}|t_{k-1}) = \Delta$. In other words, we choose the value $\overline{\Delta} > 0$, so that it solves

$$E_X(x) = \int_0^{\overline{\Delta}} x dG(x) = \Delta. \tag{25}$$

Then, in the stochastic setting, Δ is the expected monitoring intensity or the expected time interval between repetitions of the stage game. We do this in order to later perform meaningful comparisons between the deterministic and the stochastic monitoring cases. Then, $G(x)$ is a continuous differentiable distribution with support on the interval $\left(0, \overline{\Delta}\right)$, and $\overline{\Delta} > 0$ satisfies (25), i.e. during the time interval $\left(0, \overline{\Delta}\right)$, a monitoring event occurs with probability one; however, it is not known when.

It is usually assumed that the stage game is repeated at predetermined moments in time and that players discount the future according to a common discount factor, which is a convex function of time. For that reason, and without loss in generality, we assume exponential discounting.[9] When monitoring is deterministic

$$\delta^{\Delta} \equiv e^{-r\Delta}.$$

The value $r \in (0, \infty)$ denotes the discount rate. We will not make assumptions about the value of Δ and we also consider the possibility of this parameter being a random variable. Then the discount factor is no longer a deterministic but rather a stochastic function of time. The expected discount factor is denoted and equal to

$$E_X(\delta^x) \equiv \int_0^{\overline{\Delta}} e^{-rx} dG(x). \tag{26}$$

[8]As usual, an appropriate normalization will turn the payoffs of the infinitely repeated game into the same units as the stage game.

[9]The qualitative features of the results will not have changed if we have assumed hyperbolic discounting $\delta^{t_k} \equiv 1/(1 + rt_k)$ or $\delta^t \equiv 1/(1+r)^{t_k}$.

Under the expected utility hypothesis, player i's ex-ante expected payoff, denoted as $\pi_i(a)$, is the relevant element for studying the game. The stage game payoffs are independent of the monitoring intensity, i.e. independent of Δ. The effect of the time interval between actions operates through the common discount factor and the distribution of the signals.

Denote the signals space as Y, with generic element y. Depending on the stochastic process considered, the set Y changes. Typically, optimal behavior requires a partition the signal space into two types of signals; the set of signals that are interpreted as or suggest cooperative behavior, denoted as Y^+, and signals that suggest deviating behavior, Y^-. When the strategy requires the game to start with the profile a, let $p(\Delta) \equiv \Pr(y \in Y^- | \Delta, a)$ be the probability with which a bad signal is observed at the end of the period. In the case of a deviation from the equilibrium path, $q(\Delta) \equiv \Pr(y \in Y^- | \Delta, (a_i', a_{-i}))$ denotes the analogous probability.

The publicly observed history is $h^{t_k} \equiv \{y^{t_0}, y^{t_1}, ..., y^{t_{k-1}}\}$, with $h^{t_0} \equiv \varnothing$. Each player i also has a private history $h_i^{t_k} \equiv \{y^{t_0}, a_i^{t_0}, y^{t_1}, a_i^{t_1}, ..., y^{t_{k-1}}, a_i^{t_1}\}$, made up of the collection of observed public signals and the privately chosen actions. A pure strategy for player i is a mapping σ_i from the set of all possible histories $H_i \equiv \cup_{k=0}^{\infty} (A_i \times Y)^{t_k}$, into the set of pure actions A_i.

When monitoring is perfect, the signals are perfectly informative about the chosen profile but they do not provide information about the identity of the deviator.[10] In such a setting we are looking at strategy profiles $\sigma \equiv (\sigma_1, ..., \sigma_n)$ that form a *subgame perfect equilibria* (SPE Henceforth). Under imperfect public monitoring, we look at strategy profiles that are *perfect public equilibria* (PPE Henceforth).[11]

[10] This is a more restrictive notion of perfect monitoring, but does not pose restrictions on the existing folk theorems. In two player games they are equivalent.

[11] A strategy is public if it depends only on the public histories and not on the private history of player i. Given a public history, a profile of public strategies that induces a Nash equilibrium on the continuation game from that time on, is called a *perfect public equilibrium* (PPE). Moreover, if the other players $-i$ are playing public strategies, player i's best reply can only be a public strategy.

2.2.1 Deterministic Monitoring

In this case the repetitions of the stage game are deterministic, with known length Δ. Then a deterministically repeated game can be of perfect monitoring or imperfect public monitoring. We tailor this Section for the latter setting, since the former can be seen as just a particular case.

Player i's infinite sum of ex-ante normalized expected payoffs from the repeated game, under the strategy profile σ, is given by

$$v_i(\sigma) = \left(1-\delta^{\Delta}\right)\sum\nolimits_{k=1}^{\infty} \delta^{(k-1)\Delta} E_Y\left[\pi_i\left(a^{(k-1)\Delta}|\Delta,\sigma,h^{(k-1)\Delta}\right)\right].$$

Notice that the expectation is taken with respect to the full signal space Y and is conditional to the observed public history and the strategy profile σ.

With a partition of the signal space into two sets, of good and bad news signals, we can write the usual enforceability definition, extended with an equilibrium condition on the set of continuations W. That is, we restrict the continuations to elements of the set of equilibrium payoffs $V(\Delta, r)$. Following Abreu, Pearce and Stacchetti (1986, 1990) we now define an equilibrium.

Definition 8 *A profile $a \in A$ is enforceable with respect to r and Δ, on $W \subseteq V(\Delta, r)$, and $v(a, w, \Delta, r)$ is an equilibrium payoff profile, if there exists a mapping $w : Y \to W$, such that, for each player i and $a_i' \in A_i$,*

$$v_i(a, w, \Delta, r) = \left(1-\delta^{\Delta}\right)\pi_i(a) + \delta^{\Delta} E_Y\left[w_i(y)\,|a,\Delta\right],$$

and

$$a_i = \arg\max_{a_i' \in A_i}\left(1-\delta^{\Delta}\right)\pi_i\left(a_i', a_{-i}\right) + \delta^{\Delta} E_Y\left[w_i(y)\,|\left(a_i', a_{-i}\right), \Delta\right].$$

In the definition, we bundle together enforceability and self-generation. The main propose of this definition is to compare it with the analogous definition for the case where monitoring is random.

2.2.2 Random Monitoring

As mentioned before, the sequence of times $t_1, t_2, ...$ might be unknown with each value t_k, $k \geq 1$, being repeatedly drawn from some distribution.[12]

Definition 9 *We say that a repeated game is of random monitoring if the time repetitions $t_{k\geq 1}$ of the stage game are stochastic.*

As in the deterministic case, a repeated game of random monitoring can be of perfect or imperfect monitoring, depending on the informativeness of the observed signals. Random monitoring refers to the uncertainty of the time repetitions of the stage game. The difference will be clear from the context.

Since the repetitions of the stage game are not known when the game starts at time $t_0 \equiv 0$, from a strategic point of view, we should expect it to discipline the deviator. We will see that in contexts of imperfect monitoring this might not be true.

Random monitoring brings uncertainty to the time domain of the repeated game. For that reason we put together uncertainty about the public signals and about the time repetitions of the stage game. These two types of uncertainty are not independent. The reason is that the distribution of the public signals also depends on time. For example, when the public signals are Brownian the probability of an extreme event increases with the time interval Δ. For the Poisson process, the larger the time interval, the greater the number of events that are likely to be observed.[13]

Notice that the length of each time interval is independent of the length of the previous or subsequent interval. The *i.i.d.* assumption allows us to write $x = x_k \equiv t_k - t_{k-1}$, for all $k \geq 1$. Fix a k, for $t_{k-1} - t_{k-2}$, $t_{k-2} - t_{k-3}$, ..., $t_1 - t_0$ in which we can compute each

[12] We abstain here to discuss on how players are informed about the time to play the stage game. Rather we focus on studying the associated expected payoffs. However, we can think of a public signal or some sort of communication.

[13] For the special case of exponential distributed signals, the larger Δ is, the higher the likelihood that an event occurs.

discount factor independently, according to

$$E\left(\delta^{t_{k-1}}\right)=E_X\left(\delta^{(t_{k-1}-t_{k-2})+(t_{k-2}-t_{k-3})...+(t_2-t_1)+(t_1-0)}\right)=\prod_{j=1}^{k-1}E_X\left(\delta^x\right)=E_X\left(\delta^x\right)^{k-1}.$$[14]

Then, when the repetitions of the stage game are random, the normalized expected discounted stream of payoffs is given by

$$\widetilde{v}\left(\sigma\right)=\left(1-E_X\left(\delta^x\right)\right)\sum_{k=1}^{\infty}E_X\left(\delta^x\right)^{k-1}E_XE_Y\left[\delta^x\pi_i\left(a^{t_{k-1}}|x,\sigma,h^{t_{k-1}}\right)\right].$$

Notice that now, the expectation is also taken with respect to the time interval between repetitions of the stage game, which are stochastic. The idea is that, once at time t_{k-1}, the observed public signal $y^{t_{k-1}}$ at t_k, depends on the length of the time interval, which is stochastic. More precisely, the distribution of the public signals has a stochastic parameter t_k-t_{k-1}. For that reason we cannot separate discounting from the observed signal, within each period.

Consequently, we write an analogous version of Definition 8, to encompass the stochastic monitoring case.

Definition 10 *A profile $a\in A$ is enforceable with respect to r and $\Delta=E_X\left(x\right)$, on $\widetilde{W}\subseteq\widetilde{V}\left(\Delta,r\right)$, and $\widetilde{v}\left(a,w,\Delta,r\right)$ is an equilibrium payoff profile, if there exists a mapping $\widetilde{w}:Y\to\widetilde{W}$, such that, for each player i and $a_i'\in A_i$,*

$$\widetilde{v}_i\left(a,w,\Delta,r\right)=\left(1-E_X\left(\delta^x\right)\right)\pi_i\left(a\right)+E_XE_Y\left[\delta^x\left(\widetilde{w}_i\left(y\right)|a,x\right)\right],$$

and

$$a_i=\arg\max_{a_i'\in A_i}\left(1-E_X\left(\delta^x\right)\right)\pi_i\left(a_i',a_{-i}\right)+E_XE_Y\left[\delta^x\left(\widetilde{w}_i\left(y\right)|\left(a_i',a_{-i}\right),x\right)\right].$$

Notice now, that the value of the repeated game after the first play of the game is an expectation not only over the distribution of the public signals but also over the distribution of the repetitions of the stage game. We cannot separate the discounting from the continuation value. This is the main difference with the canonical deterministic representation.

[14] We can obtained the same result if, instead, we compute the recursive interacted conditional expectation $E_{t_1}\left(E_{t_2}\left(...E_{t_{k-1}}\left(E_{t_k}\left(\delta^{t_k-t_{k-1}}|t_{k-1}\right)|t_{k-2}\right)...|t_1\right)|t_0\right)=E\left(\delta^x\right)^k$, for $t_k=t_{k-1}+x$, with the expectation taken w.r.t. the i.i.d. random variables x.

Definitions 8 and 10, in this Section, are general enough to accommodate both perfect and imperfect monitoring in deterministic and stochastic environments.

We now raise the following questions about random monitoring in repeated games: Is it good for the provision of incentives? What are the effects on the distribution of the public signals? Does the set of continuations and equilibrium payoffs differ, i.e. $\widetilde{V}(\Delta, r) \neq V(\Delta, r)$? Depending on whether we consider perfect or imperfect public monitoring, the following Sections attempt to provide an answer to these questions.

2.3 Perfect Monitoring

Under perfect monitoring with random repetitions of the stage game, due to the time uncertainty, the crucial aspect is the discount factor. The particular case of perfect monitoring, is also covered by Definitions 8 and 10. When the stage game is repeated at predetermined and equally spaced moments in time, players discount the future according to δ^{Δ}. On the other hand, when the repetitions are not predetermined, the discount factor to apply is $E_X(\delta^x)$. In the previous section we have seen the difference between these discount factors.

Before going further, we shall add some more notation while we review some important concepts in repeated games. Denote $\pi_i(a)$ as an arbitrary player's i stage game payoff on the equilibrium path, and $\pi_i(a_i', a_{-i})$ as the player's i payoff of the most profitable stage game deviation from a. Let the pair, $\pi_i(a)$ and $\pi_i(a_i', a_{-i})$, form the most attractive deviation along the equilibrium path, for all $i \in N$ and all $a_i' \in A_i$.

In line with Definition 8, if for each action profile a there are credible continuation promises $w(a) \in V(\Delta, r)$ such that for each player $i \in N$, and all $a_i' \in A_i$,

$$v_i = \left(1 - \delta^{\Delta}\right) \pi_i(a) + \delta^{\Delta} w_i(a) \geq \left(1 - \delta^{\Delta}\right) \pi_i\left(a_i', a_{-i}\right) + \delta^{\Delta} w_i\left(a_i', a_{-i}\right), \tag{27}$$

then the profile a is enforceable on $V(\Delta, r)$ and the payoff profile $v \in V(\Delta, r)$ can be sustained as a pure-strategy SPE.[15]

The same concept can be generalized with random monitoring, by simply substituting δ^{Δ} for $E_X(\delta^x)$ in (27). With a slight abuse of notation, let $\widetilde{V}(\Delta, r) \subset \mathbb{R}^n$ ($V(\Delta, r) \subset \mathbb{R}^n$)

[15] As in Definition 8, we directly assume that any value v and $w(a)$ belongs to $V(\Delta, r)$, for that reason we do not mention the "largest self-generating property" of the set of subgame-perfect equilibria payoffs.

denote the set of pure-strategy SPE payoffs under stochastic (deterministic) monitoring, with the generic element $\widetilde{v}$ (v).

We now present our first result. Fix a monitoring intensity $\Delta = E_X(x)$, in deterministic or random terms. Suppose we have a particular pure-strategy SPE payoff profile $v \in \mathbb{R}^n$ that can be sustained both by monitoring randomly, i.e. $v \in \widetilde{V}(\Delta, r)$, and by deterministic monitoring, i.e. $v \in V(\Delta, r)$. Then we ask; to sustain v as an equilibrium payoff, which monitoring technology allows a higher discount rate?

Proposition 11 *Given a monitoring intensity $\Delta = E_X(x)$, to sustain a particular pure-strategy SPE payoff profile v, deterministic monitoring requires a lower discount rate than random monitoring, if $\delta^x \in (0,1)$ is strictly convex in x.*

The result states that the same payoff v can be achieved with higher impatience levels when the repetitions of the stage game are stochastic. Random monitoring enlarges the spectrum of impatience where a given equilibrium payoff can be sustained. This is a striking result in the theory of repeated games.

Intuitively, uncertainty in the time domain disciplines potential deviating behavior. Now, more impatient players prefer to stay on the equilibrium path, rather than deviating and obtaining an expected gain of uncertain duration.

Moreover, we stress that it is commonly assumed that players discount the future in a convex way, according to either an exponential or a hyperbolic discount function. For that reason, it is hard to think of situations where the convex condition of Proposition 11 fails. Under perfect monitoring, only when δ^x is not strictly convex or when the distribution of the random time is degenerate, does the result not apply.

While, for the sake of simplicity, we focus on pure strategies, the results of Proposition 11 can be immediately extended for behavioral/mixed or correlated strategies.

Example 12 *Suppose that $X \sim U\left(0, \overline{\Delta}\right)$ is uniformly distributed with $\overline{\Delta} = 2\Delta$, implying that $E_X(x) = \Delta$ as required in Proposition 11. In the deterministic monitoring case, a* *folk theorem in pure-strategies for the prisoners' dilemma of Table 1 below,*[16] *with* $\pi' = 3$, $\pi = 2$ *and* $\Delta = 1$, *can be obtained providing that players' discount rate* $r \leq 0.4055$. *For the same expected monitoring frequency, the same folk theorem can be obtained with random monitoring, with a higher discount rate* $\widetilde{r} \leq 0.4371$. *Similarly, if we allow players to use public correlation, we can obtain a folk theorem for discount rates* $\widetilde{r} \leq 1.4107$, *rather than* $r \leq 1.0986$.[17]

Proposition 11 and the numerical example show that a folk theorem under random monitoring can be sustained with a higher discount rate.

We can also reinterpret Proposition 11, by saying that there are payoffs which cannot be sustained under deterministic monitoring, but which can be sustained as pure-strategy SPE if the repetitions of the stage game are unknown to the players. The following result formalizes this intuition.

Corollary 13 *When the convex condition of Proposition 11 holds, given a discount rate* r *and an expected monitoring intensity* Δ, *there are pure-strategy SPE payoffs that can be sustained only under random monitoring, while the converse is not true.*

The result tells us that, depending on r and Δ, the set of SPE obtained with deterministic monitoring is a subset of the set of SPE obtained with stochastic monitoring, that is $V(\Delta, r) \subseteq \widetilde{V}(\Delta, r)$. To see this more clearly, consider the following example, which is a continuation of Example 12 presented above.

Example 14 *Using Corollary 13, we can reinterpret the above numerical example. Suppose* $r = 0.42 > 0.4055$, *with deterministic monitoring we are able to sustain a great number of equilibria but not a folk theorem. However, with stochastic monitoring, since* $0.42 \in (0, 0.4371]$, *we can sustain any feasible and weakly individual rational payoff profile. With*

[16] *We say that a folk theorem exists if any feasible and strictly individual rational payoff can be supported by a strategy that forms a subgame-perfect equilibrium of the infinitely repeated game. For folk theorems with discounting in infinitely repeated games, see the classical papers of Friedman (1971) and Fudenberg and Maskin (1986).*

[17] *This example is adapted from Mailath and Samuelson (2006, sections 2.5.3 and 2.5.6).*

public correlation, if $r = 1.3 > 1.0986$*, we simply sustain the infinite repetition of the stage game Nash equilibrium.*[18] *However, when monitoring is random, we obtain a full folk theorem since* $1.3 \in (0, 1.4107]$.

It is worth noticing that, when the monitoring is random, the players' true discount factor does not change. However, under the expected utility hypothesis, their decisions are based on a larger discount factor (the expected discount factor). This reflects the intuition behind these results. The uncertainty about the moment when the stage game is repeated brings uncertainty about the value of gains that a potential deviator may contemplate; in expected terms they get smaller. The continuation value of the game becomes more important. In some sense, the effect is similar, as if the players had become "more patient". We call this effect, *the expected discount factor effect*.

Definition 15 *For fixed* r *and* $\Delta = E_X(x)$*, we say that there is an expected discount factor effect when* $E_X(\delta^x) > \delta^\Delta$*.*

One aspect of the expected discount factor effect, that we did not explore in great detail in this section, is that it tends to be stronger for lower frequencies of monitoring, i.e. larger values of Δ. This fact will be useful to interpret some of the results in the following Sections.

In short, random monitoring strictly improves on deterministic monitoring in contexts where signals are perfectly informative; this is the conclusion of the present Section.

2.4 Imperfect Public Monitoring

In the previous Section, we have seen that stochastic monitoring favors the provision of incentives in contexts of perfect monitoring. This was due to the expected discount factor effect, and holds for any feasible expected monitoring intensity Δ. However, the result does not generalize in the same straightforward sense when monitoring is imperfect. The difference is that the informativeness of the public signals might be adversely affected by the uncertainty on the repetitions of the stage game for some monitoring intensities.

[18] *Stahl (1991) discusses the discontinuity on the set of subgame-perfect equilibrium payoffs with public correlation in detail.*

	C	D
C	π, π	$-(\pi' - \pi), \pi'$
D	$\pi', -(\pi' - \pi)$	$0, 0$

Table 1: The Prisoners' Dilemma Stage Game Payoffs.

In order to study the effects of random monitoring in repeated games with public signals, for simplicity, we present our results for a classical prisoners' dilemma, whose payoffs are shown in Table 1.[19]

We assume $\pi' > \pi > 0$ and $2\pi > \pi'$, so that defection is a dominat strategy for both players. The minimax value of the game coincides with the stage game Nash payoffs and equals 0 for both players.

Our focus is on the *strongly symmetric equilibrium (SSE* henceforth); where the same action is chosen by both players after every public history. That is, we want to sustain the infinite repetition of the profile (C, C).

The difference between perfect monitoring (studied in the previous section) and an imperfect public monitoring is that, at each moment in time $t_1, t_2, ...$, players simultaneously adjust their actions, but also observe a public signal which provides noisy information about the action profile resulting from each player's private choices.

2.4.1 The Best Strongly Symmetric equilibrium

The expression that gives the value of best SSE of the prisoners' dilemma of Table 1 is well known. However, in order to accommodate the random monitoring case, we need to consider some specificities. This is the goal of this section.

In the canonical deterministic monitoring setup, of length Δ, the probability of observing a bad signal when the profile (C, C) has been chosen, is given by

$$p(\Delta) \equiv \int_{Y^-} f(y, a)\, dy, \tag{28}$$

[19] We restrict our analysis to a simple setting. We do this in order to concentrate on random monitoring and associated informational effects, without extra complexities. Such a restriction does not diminish in any way the point this paper wants to address. The analysis extends to other games.

and in the case of a deviation, i.e. the profiles (C, D) or (D, C),

$$q(\Delta) \equiv \int_{Y^-} f\left(y, \left(a'_i, a_{-i}\right)\right) dy. \tag{29}$$

Where, $f(y, a)$ and $f(y, (a'_i, a_{-i}))$ are the densities of the distribution of public signals in the cases of mutual cooperation and unilateral defection, respectively.

An observation of a realized signal, inside the support Y^-, is interpreted as a bad signal. For example, when signals are Brownian; with one threshold $Y^- \equiv (-\infty, b]$, and with two symmetric thresholds: $Y^- \equiv (-\infty, -b] \cup [b, \infty)$ when the extreme realizations are bad news, and $Y^- \equiv [-b, b]$ when small magnitude observations are bad news.[20] When the signals follow a Poisson process; $Y^- \equiv [0, x]$ in the bad news case, and $Y^- \equiv [x, \infty]$ in the good news case (with $x = \Delta$ if monitoring is deterministic).

Clearly, we should expect $q(\Delta) > p(\Delta)$, meaning that a deviation has an higher probability of generating a "bad signal".[21]

When monitoring is stochastic, the realizations of the public signal are not independent of the random time at which they are observed. Only when a random monitoring event occurs, do players observe the realized signals and collect their payoffs, which are discounted to the present from that moment in time. For that reason, we cannot separate discounting from the distribution of the public signals. Then, the punishment "probabilities" have to be adapted, where

$$\widetilde{p}_\delta(\Delta) \equiv \int_0^{\overline{\Delta}} \int_{Y^-} e^{-rx} g(x) f(y, a) \, dy dx, \tag{30}$$

and

$$\widetilde{q}_\delta(\Delta) \equiv \int_0^{\overline{\Delta}} \int_{Y^-} e^{-rx} g(x) f\left(y, \left(a'_i, a_{-i}\right)\right) dy dx, \tag{31}$$

denote "expected discounted punishment probabilities". Here $g(x)$ is the density function of the random repetitions of the stage game, and $e^{-rx} \equiv \delta^x$ is the discount factor that

[20] Fudenberg and Levine (2007) show that when defective behavior impacts on the volatility of the process, the optimal monitoring technology uses two thresholds. With no impact on the drift, these thresholds are symmetric.

[21] For more demanding informational conditions on the public signals, see Fudenberg, Levine and Maskin (1994), and Fudenberg and Levine (1994).

follows the same probability law. The equal Δ in (28), (29), (30) and (31) emphasizes the fact that the monitoring intensities are the same in expected terms.

Even though discounting and signals are convolved, as in (30) and (31), to find the expression of the value of the best SSE, we are still able to apply the recursive dynamic programming methods developed by Abreu, Pearce and Staccetti (1986, 1990).

In the deterministic case, the expected continuation value of the game is some convex combination between the expected normalized payoff when play starts with the observation of a good signal, and the expected normalized payoff when play starts with an observation of a bad signal, i.e. $(1 - p(\Delta))\,\overline{v} + p(\Delta)\,\underline{v}$. However, when monitoring is stochastic, the "expected continuation value" cannot be separated from discounting, i.e. $(1 - \widetilde{p}_\delta(\Delta))\,\widetilde{v} + \widetilde{p}_\delta(\Delta)\,\underline{\widetilde{v}}$. Then it makes more sense to speak about an "expected discounted continuation value", as in Definition 10 of Section 2.2.

Lemma 16 *(i) Under deterministic monitoring of frequency Δ, the value of the best SSE is given by*

$$\overline{v}(\Delta) = \pi - \frac{p(\Delta)}{q(\Delta) - p(\Delta)}\left(\pi' - \pi\right), \tag{32}$$

with

$$\frac{1}{\delta^{\Delta}} - \frac{q(\Delta)\,\pi - p(\Delta)\,\pi'}{\pi' - \pi} = \alpha(\Delta), \tag{33}$$

and while $\alpha(\Delta) \leq 1$.

(ii) Under random monitoring of expected frequency Δ, the value of the best SSE is given by

$$\widetilde{v}(\Delta) = \pi - \frac{\widetilde{p}_\delta(\Delta)}{\widetilde{q}_\delta(\Delta) - \widetilde{p}_\delta(\Delta)}\left(\pi' - \pi\right), \tag{34}$$

with

$$\frac{1}{E_X(\delta^x)} - \frac{\widetilde{q}_\delta(\Delta)\,\pi - \widetilde{p}_\delta(\Delta)\,\pi'}{E_X(\delta^x)\,(\pi' - \pi)} = \widetilde{\alpha}(\Delta), \tag{35}$$

and while $\widetilde{\alpha}(\Delta) \leq 1$.

Expressions (32) and (34) characterize the value of the best SSE payoffs, under deterministic and random monitoring, respectively. Notice that, when monitoring is stochastic, the expression for the value of the best SSE payoff in (34) depends directly on the discount rate r. This is a consequence of the fact that discounting and realized signals are not independent.

We should note that it is natural to interpret $\alpha(\Delta)$ and $\widetilde{\alpha}(\Delta)$ as probabilities, taking values on the interval $[0,1]$. Such is the case when y takes a continuum of values, as in the case of Brownian signals, where optimality requires $\alpha(\Delta) = \widetilde{\alpha}(\Delta) = 1$. However, the Poisson signals are discrete, in particular when signals follow an exponential distribution, so at any given moment in time players only see if an "event" has occurred or not. For this reason, the left-hand side of (33) and (35) may take negative values when we force optimality. Nonetheless, with $\alpha(\Delta)$ or $\widetilde{\alpha}(\Delta)$ taking negative values, we always have $\underline{\widetilde{v}}(\Delta) \in [0, E_X(\delta^x)\widetilde{v}(\Delta)] \subset [0, \widetilde{v}(\Delta))$ and $\underline{v}(\Delta) \in \left[0, \delta^\Delta v(\Delta)\right] \subset [0, v(\Delta))$, as it should. For this reason, and without loss in generality, we ignore this technical issue in order for Lemma 16 to be general enough to harbour the Brownian and the Poisson cases.[22] In fact, when $\alpha(\Delta)$ or $\widetilde{\alpha}(\Delta)$ go below zero, this means that one punishment period is too severe. An intertemporal transfer of value is needed to compensate for this fact.

2.4.2 Efficiency Gains with Random Monitoring

In Section 2.3, we have seen that for any feasible monitoring intensity, perfect random monitoring always allows efficiency gains with respect to perfect deterministic monitoring. Even when players' true discount rate remain unchanged, their decisions were based on a small discount rate. We call this the expected discount factor effect, which was the key for the obtained results.

Under random public monitoring, the expected discount factor effect is still present and with the same intensity. The difference is that, now, the uncertainty of the repetitions of the stage game may also adversely affect the distribution of the public signals. For this reason, it is not clear if public random monitoring is better than the canonical deterministic case.

[22] We can go around this issue by correlating the punishment decision. However, this case is not free of technical issues either. In particular, optimality requires players to ignore a bad signal with a probability that converges to 1 as $\Delta \to 0$.

The following result establishes conditions under which public random monitoring allows efficiency gains on the value of the best SSE. The superscript "$*$" denotes optimal behavior.

Proposition 17 *Given a monitoring intensity* $\Delta = E_X(x)$ *and a common discount rate* r*, in the infinitely repeated prisoners' dilemma described above, the value of the best SSE* $\widetilde{v}^*(\Delta)$*, is larger than the value of the best SSE* $\overline{v}^*(\Delta)$*, in either of the following cases;*

(i) When $\widetilde{\alpha}^*(\Delta) \leq 1$*,* $\alpha^*(\Delta) \leq 1$*, and*

$$\widetilde{q}_{\delta}^*(\Delta) / \widetilde{p}_{\delta}^*(\Delta) \geq q^*(\Delta) / p^*(\Delta). \tag{36}$$

(ii) When $\widetilde{\alpha}^*(\Delta) \leq 1$ *and* $\alpha^*(\Delta) \geq 1$*.*

In part (i), random monitoring leads to an efficiency gain equal to $\widetilde{v}^*(\Delta) - \overline{v}^*(\Delta)$, while in scenario (ii), the efficient gain is equal to $\widetilde{v}^*(\Delta)$. In the former, the profile (C, C) can be enforced with both random and deterministic monitoring. Then, it is condition (36) that determines when random monitoring is superior in payoff terms. In the latter scenario, the profile (C, C) can only be enforced with random monitoring.

Outside the cases stated in Proposition 17, either deterministic monitoring is more efficient or no monitoring technology can improve over the static Nash.

Notice that $\widetilde{v}^*(\Delta)$ and $\overline{v}^*(\Delta)$ are also upper bounds on the set of SSE payoffs. When $\widetilde{v}^*(\Delta) \geq \overline{v}^*(\Delta)$, we can also say that the set of SSE payoffs of an infinitely repeated game played at random moments in time is larger.

The consistent improvement of stochastic monitoring over deterministic monitoring shown in Section 2.3, cannot be generalized in the same way under public monitoring. Proposition 17, establishes conditions where efficiency gains are possible, for fixed Δ and r. It does not say that random monitoring is better for all feasible Δ.

The expected discount factor effect was the key factor for the strong results of Section 2.3. Condition (36) is also positively affected by the expected discount factor effect. However, there is another effect associated with random monitoring which might be positive or negative, i.e. *the informational effect* on the distribution of the public signals.

In order to separate the expected discount factor effect from the informational effect in condition (36), we need to compute the undiscounted punishment probabilities under random monitoring, i.e.

$$\widetilde{p}(\Delta) \equiv \int_0^{\overline{\Delta}} \int_{Y^-} g(x) f(y, a) \, dy dx,$$

and

$$\widetilde{q}(\Delta) \equiv \int_0^{\overline{\Delta}} \int_{Y^-} g(x) f\left(y, \left(a_i', a_{-i}\right)\right) dy dx,$$

respectively, in the cooperative and unilateral defection cases. These probabilities take into account the amount of uncertainty about the repetitions of the stage game, but do not incorporate discounting. So we have the pure informational effect.[23]

Definition 18 *We say that random monitoring has a positive informational effect when*

$$\widetilde{q}^*(\Delta) / \widetilde{p}^*(\Delta) > q^*(\Delta) / p^*(\Delta). \tag{37}$$

Otherwise, the informational effect is negative or adverse.

We know that $q^*(\Delta) > p^*(\Delta)$ and $\widetilde{q}^*(\Delta) > \widetilde{p}^*(\Delta)$ for all Δ. The definition says, that if, by randomly repeating the stage game, the difference between $\widetilde{q}^*(\Delta)$ and $\widetilde{p}^*(\Delta)$, decreases with respect to the difference between $q^*(\Delta)$ and $p^*(\Delta)$, then signals become less informative for a given monitoring intensity Δ.

It is worth noticing that random monitoring tends to move $\widetilde{q}^*(\Delta)$ and $\widetilde{p}^*(\Delta)$ in the same direction when signals are Poisson; in the bad news case $\widetilde{q}^*(\Delta) < q^*(\Delta)$ and $\widetilde{p}^*(\Delta) < p^*(\Delta)$, because the punishment probabilities are concave in the Δ domain; while in the good news case, we observe $\widetilde{q}^*(\Delta) > q^*(\Delta)$ and $\widetilde{p}^*(\Delta) > p^*(\Delta)$, due to the convexity in Δ. However, this fact is not crucial for the relation between the values of the SSE in part (i). Moreover, under Brownian signals this might not always be the case.[24]

Corollary 19 *When the ratio (37) is satisfied, then (36) holds.*

[23] The product of the disaggregated expected discount factor and informational effect is not equal to the total effects, because the convolution with discounting generates cross effects. However, the approximation is sufficiently close, since cross effect are typically of infinitesimal magnitude.

[24] When signals are Brownian, typically $\widetilde{q}^*(\Delta) > q^*(\Delta)$ and $\widetilde{p}^*(\Delta) > p^*(\Delta)$, but for some monitoring intensities these relations might not be satisfied.

Providing that $\widetilde{\alpha}^*(\Delta) \leq 1$ and $\alpha^*(\Delta) \leq 1$ are satisfied, when both the expected discount factor effect and the informational effect play in the same direction, we must have a payoff improvement.

However, condition (37) might fail, but we are still able to improve with random monitoring. It all depends on the strength of both effects, which in this case, play in opposite directions. For this reason, the statement of Corollary 19 is a sufficient condition for random monitoring to improve on deterministic monitoring.

For example, in the Poisson good news case of Abreu, Milgrom and Pearce (1991), in the Osório-Costa (2008) model and in the Fudenberg and Levine (2007) volatility sensitive model, the inequality (37) is never satisfied, because the punishment probabilities are convex in Δ. If payoff improvements are possible, either in the sense of Part (i) or Part (ii) of Proposition 17, this is due exclusively to the expected discount factor effect, which becomes stronger as Δ increases.[25]

With respect to the convexity of punishment probabilities, we typically have $p^*(\Delta) < \widetilde{p}^*(\Delta)$ and $q^*(\Delta) < \widetilde{q}^*(\Delta)$. The expectation causes an adverse effect since, under optimal behavior, $p^*(\Delta)$ is relatively more convex than $q^*(\Delta)$. For this reason $\widetilde{q}^*(\Delta)/\widetilde{p}^*(\Delta)$ decreases with respect to $q^*(\Delta)/p^*(\Delta)$, because $\widetilde{p}^*(\Delta)$ increases relatively more than $\widetilde{q}^*(\Delta)$.

However, when the events represent bad news, for sufficiently large μ/β and Δ, condition (37) can be satisfied and we observe a positive informational effect. This is due to the decreasing concavity of the punishment probabilities with Δ. In this case $\widetilde{p}^*(\Delta) < p^*(\Delta)$ and $\widetilde{q}^*(\Delta) < q^*(\Delta)$. While for small Δ, $\widetilde{p}^*(\Delta)$ decreases relatively less than $\widetilde{q}^*(\Delta)$, when Δ is large, we might observe that $\widetilde{p}^*(\Delta)$ decreases relatively more than $\widetilde{q}^*(\Delta)$, causing the positive informational effect.

The preceding discussion provides the intuition for the informational effect under random public monitoring, which is crucial to understand when improvements in the sense of Part (i) of Proposition 17 are possible. However, there is another important way in which random

[25] An increase in r or Δ increases the relative convexity between $E_X\left(e^{-rx}\right)$ and $e^{-r\Delta}$, enhancing the expected discount factor effect.

monitoring improves on deterministic monitoring, i.e. in the sense of Part (ii) of Proposition 17. This case occurs for monitoring intensities where $\underline{\widetilde{v}}^{*}(\Delta) \geq 0 \geq \underline{v}^{*}(\Delta)$, i.e.

$$\frac{\widetilde{q}_{\delta}^{*}(\Delta)\pi - \widetilde{p}_{\delta}^{*}(\Delta)\pi'}{1 - E_X(\delta^x)} \geq \pi' - \pi \geq \delta^{\Delta}\frac{q^{*}(\Delta)\pi - p^{*}(\Delta)\pi'}{1 - \delta^{\Delta}}. \tag{38}$$

This condition tends to hold in some cases, typically for large Δ, i.e. when the expected discount effect is stronger. However, such a conclusion needs more careful consideration.

We can approximate the inequality (38) in a similar way to that done with the ratio (37). The goal is to separate the discount factor effect from the informational effect, in order to get a more intuitive condition. After some algebra, the chain of inequalities (38) is approximately,

$$\frac{E_X(\delta^x)}{1 - E_X(\delta^x)}\left(\widetilde{q}^{*}(\Delta)\pi - \widetilde{p}^{*}(\Delta)\pi'\right) \gtrsim \pi' - \pi \geq \frac{\delta^{\Delta}}{1 - \delta^{\Delta}}\left(q^{*}(\Delta)\pi - p^{*}(\Delta)\pi'\right), \tag{39}$$

since $E_X(\delta^x)\,\widetilde{q}^{*}(\Delta) \approx \widetilde{q}_{\delta}^{*}(\Delta)$ and $E_X(\delta^x)\,\widetilde{p}^{*}(\Delta) \approx \widetilde{p}_{\delta}^{*}(\Delta)$.[26]

There are two relevant factors that play a role when we consider random monitoring improvements in the sense of Part (ii) of Proposition 17. First, the expected discount factor effect, always improves the LHS of the inequality (39), with respect to the RHS. Second, the relative magnitude of the difference between $\widetilde{q}^{*}(\Delta)\pi - \widetilde{p}^{*}(\Delta)\pi'$ and $q^{*}(\Delta)\pi - p^{*}(\Delta)\pi'$, which has implicit a relative informational effect. Independently of how the punishment probabilities change with Δ, we must have $\widetilde{q}^{*}(\Delta)\pi - \widetilde{p}^{*}(\Delta)\pi' > 0$. This is a necessary but not a sufficient condition, for random monitoring to improve in the sense of Part (ii) of Proposition 17. Otherwise, the LHS of (39) would never be above $\pi' - \pi > 0$. However, such might not be sufficient, because the relative increasing effect of $E_X(\delta^x)/(1 - E_X(\delta^x))$ with Δ might not be strong enough to compensate for the decrease in $\widetilde{q}^{*}(\Delta)\pi - \widetilde{p}^{*}(\Delta)\pi'$ when Δ gets large.

Since the expected discount factor effect always improves in absolute and relative terms, for Part (ii) of Proposition 17 to hold for small Δ, it is because $\widetilde{q}^{*}(\Delta)\pi - \widetilde{p}^{*}(\Delta)\pi'$ is larger than $q^{*}(\Delta)\pi - p^{*}(\Delta)\pi'$ or they do not differ substantially, in order for the expected discount factor effect to operate. For large Δ the expected discount factor effect is the main driving force.

[26] The cross effects are typically of infinitesimal magnitude.

In short, improvements of the kind of Part (ii) of Proposition 17, require some degree of informativeness of the random monitoring signals with respect to deterministic monitoring signals. A loss of informativeness might occur with random monitor, but it should not be too large. We conclude that improvements in the sense of Part (ii) are not so informationally demanding as improvements in the sense of Part (i) of Proposition 17.

Although we have tried to provide a general treatment, each process and each model has its own specificities, a specialized treatment might be required for a more detailed understanding.

While the Poisson model is still tractable, the Brownian case is not. For this reason, in the next Section, through a series of examples, we will illustrate some of the issues discussed in the present section.

2.5 Numerical Examples - The Prisoners' Dilemma

In this section we illustrate, using numerical examples, potential efficiency gains due to random monitoring. For the case when monitoring is imperfect but public, we study when efficiency gains in the models of frequent monitoring, where payoffs above the static Nash are possible in the limit; the Abreu, Milgrom and Pearce (1991) bad news case, the Fudenberg and Levine (2007) model, where deviations change the volatility of the process, and the Osório-Costa (2008) model, where different payoff profiles lead to different initial conditions. The main concern of these papers is on the limit case, i.e. $\Delta \to 0$. Here, we look at how these models behave for more general monitoring frequencies.

To illustrate this;

(i) We consider that at each moment in time $t_0, t_1, t_2, ...$, players repeatedly play the prisoners' dilemma stage game of Table 1, with $\pi' = 3$ and $\pi = 2$.

(ii)The minimax value of the game coincides with the stage game Nash payoffs and equals 0 for both players. The full efficient SSE payoff equals 2 for each player.

(iii) We assume that $x \sim U\left(0, \overline{\Delta}\right)$ with $\overline{\Delta} = 2\Delta$, implying that $E_X\left(x\right) = \Delta$. The uniform distribution is interesting, not only for its simplicity, but also because it maximizes the entropy of random monitoring for distribution with bounded support.

(iv) We set the discount rate, $r = 0.1$.

2.5.1 Abreu, Milgrom and Pearce (1991) - Bad News Case

In this section we replicate the work of Abreu, Milgrom and Pearce (1991) for monitoring intensities other than the limit case. For simplicity, we consider that at the end of the time interval of length Δ, players can only observe one of two possible signals; either an event has occurred, or it has not. Consequently, signals are exponentially distributed.[27]

If an event is publicly observed, the relation enters the punishment stage, returning from it after having observed some public signal.[28]

In the bad news model, when both players cooperate, events arrive at rate β, while if there is a unilateral deviation "bad news" arrives with a higher intensity $\mu > \beta > 0$.

In this case, the value of the best SSE payoff improves monotonically with the monitoring intensity, converging to some value above the static Nash as reported by Abreu, Milgrom and Pearce.[29] This fact is illustrated in Figure 6.

Independently of the monitoring frequency, the smaller the ratio β/μ is, the larger the payoffs that can the achieved. Figure 6 shows that when β/μ is small, we observe that random monitoring is superior to deterministic monitoring, in some interval of monitoring frequencies, in the sense of Part (i) and (ii) of Proposition 17. While, when β/μ takes larger values, deterministic monitoring is superior for all Δ.

Notice that random monitoring can achieve larger payoffs for all feasible monitoring intensities if we let $\beta/\mu \rightarrow 0$.

When monitoring is random, the value of the best SSE varies with the discount rate r.[30] However, the discount rate has a more important effect on the measure of the set of monitoring frequencies that sustain payoffs above the static Nash. A larger value of r,

[27] This allows us to avoid the integer problem of a Poisson process, associated with the number of events that have occurred in the time interval Δ.

[28] The resource to public correlated signal, that recommends players to stay in the punishment stage with some probability, instead of ignoring a bad signal with some probability, is chosen to make the enforceability condition binding. To be general enough, we decided to correlate on the punishment stage, allowing for intertemporal transfers within that stage.

[29] Osório-Costa (2008) reports a similar monotonic result when the public signals are Brownian.

[30] Recall that r directly enters in (34) through the stochastic discount factor.

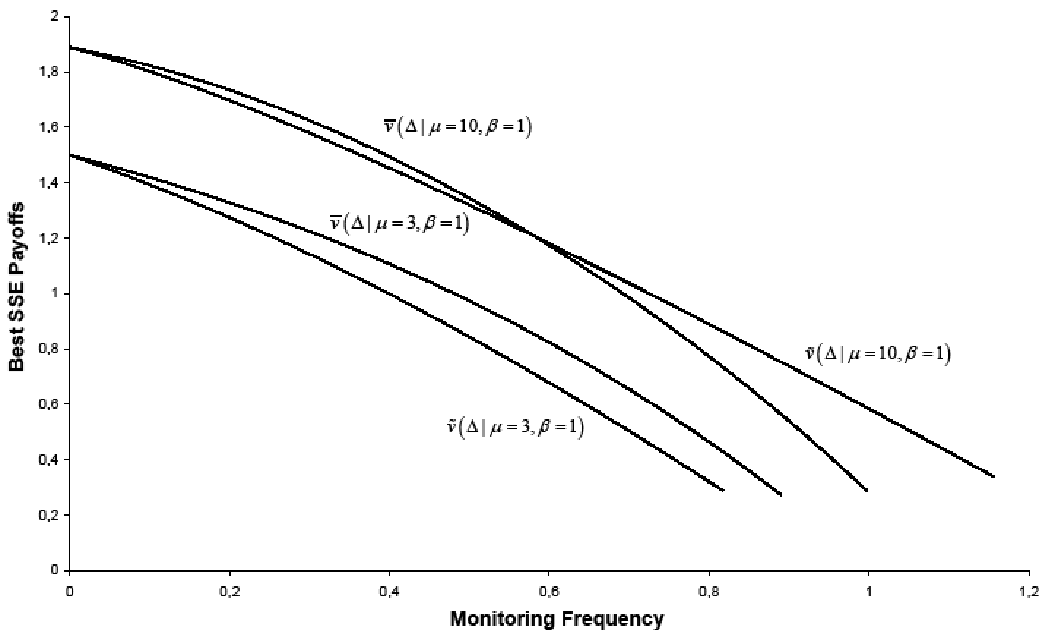

Figure 6: Random monitoring effect on the SSE payoff in the Abreu, Milgrom and Pearce (1991) bad news model.

reduces this set. This is true for random and deterministic monitoring.

Due to space constraints, we do not go through the "good news" model of Abreu, Milgrom and Pearce in detail. In this case, it is the lack of observed movements in the public process that are interpreted as "bad signals".[31] This is a degenerating type model, in the sense that at the limit, no equilibrium, other than the infinite repetition of the static Nash, can be sustained. However, away from the limit, it is possible to obtain improvement in the sense of Part (ii) of Proposition 17, for β/μ sufficiently large, enlarging the spectrum of monitoring intensities with payoffs above the static Nash.

2.5.2 Fudenberg and Levine (2007) - Volatility Sensitive Model

When a defection changes the volatility parameter of the diffusion process, it is possible to obtain non-trivial payoffs at the limit. Fudenberg and Levine (2007) show that when a deviation causes a significant increment on σ, full efficient results can be obtained at the limit. In this case, extreme values of the process suggest defective behavior. However, when a deviation decreases the uncertainty parameter, limit payoffs above the static Nash, but not fully efficient, are possible. Small realizations of the process are bad news. Figure 7 illustrates these two cases.

Even though random and deterministic monitoring converge to the same limit, for any parameterization of the model, random monitoring cannot improve in any sense of Proposition 17. The expected discount factor effect cannot compensate the informational loss of the public signals when monitoring is random.

The asymptotic efficient result obtained by Fudenberg and Levine is similar in shape to the one presented in Osório-Costa (2008). However, their intuition is different, as we shall see.

Fudenberg and Levine (2007), and Sannikov and Skrzypacz (2007), also explore frequent monitoring when players' actions exclusively change the drift of the process. Since different profiles of actions have the same associated initial condition, the set of SSE degenerates,

[31] Where μ and $\beta > \mu > 0$ denote the intensity of good news arrival, when there is a unilateral deviation and when both players cooperate, respectively.

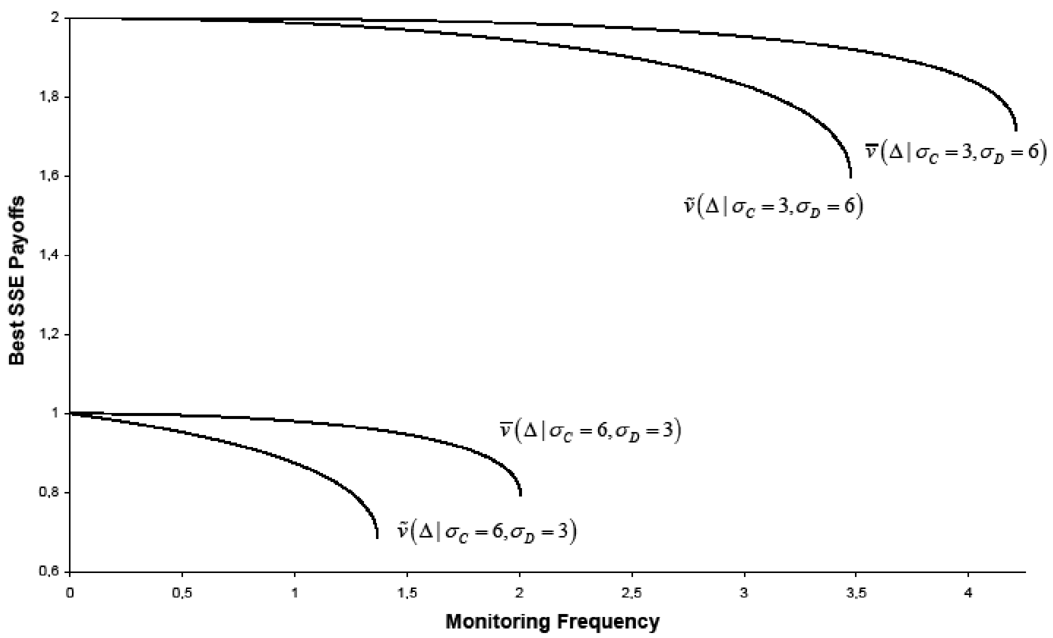

Figure 7: Random monitoring effect on the SSE payoff of the Fudenberg and Levine (2007) model. Defection increases the volatility (upper pair of curves). Defection decreases the volatility (lower pair of curves).

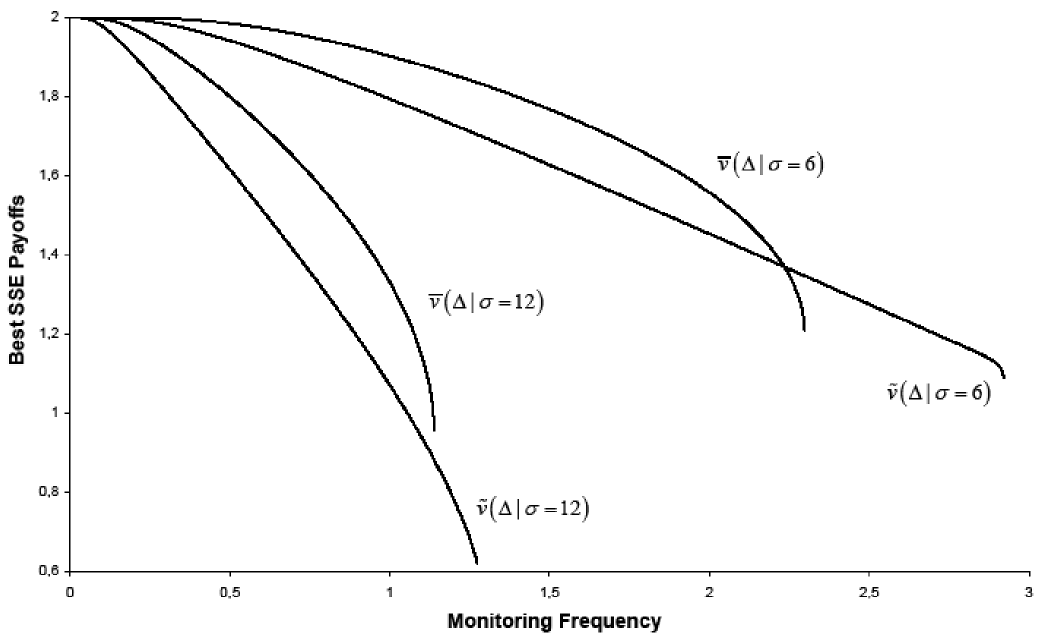

Figure 8: Random monitoring effect in the Osório-Costa (2008) model.

for high monitoring intensities.

Away from the limit, random monitoring cannot improve in any sense of Proposition 17.

2.5.3 Osório-Costa (2008) - Distinct Initial Conditions Model

In this setting, low observations of the public signal are interpreted as a "bad signals". There is a cut-off point that separates "good" from "bad" signals. Osório-Costa (2008) provides a detailed characterization of such a decision rule. In this model, different unknown profiles of actions, chosen at the beginning of each period, give rise to different initial conditions. Efficient results are obtained at the limit and payoffs improve monotonically with the monitoring intensity. The monotonicity is similar in shape to the one described above for the Abreu, Milgrom and Pearce (1991) bad news model. (See Figures 6 and 8)

Independently of the monitoring frequency, the smaller the uncertainty parameter σ is, the larger the payoffs that can the achieved. Random monitoring is superior according to Proposition 17 Part (i), in a given interval, that enlarges as σ increases and vanishes for a sufficiently large σ. Figure 8 shows improvements of this type when $\sigma = 6$ but not when $\sigma = 12$. Improvements in the sense of Part (ii) of Proposition 17 are also possible, as we can see in Figure 8, both for $\sigma = 6$ and 12, enlarging the spectrum of monitoring intensities that can sustain non trivial equilibria.[32] However, when σ gets larger, deterministic monitoring becomes the most efficient monitoring technology for all feasible Δ.

When there is an interval of monitoring intensities for which random monitoring is superior in any sense, that interval does not vanish for varying r. However, its measure and location does. Clearly, low values of r have positive effects on the payoffs.

2.6 Final Comments

In this paper we study a dynamic version of stochastic monitoring in repeated games with moral hazard. Players are uncertain about the moment in time when they will be the object of monitoring or, in other words, they are uncertain about when the next stage of the repeated game is going to be played.

We found that perfect random monitoring is always better than the classical deterministic repetition approach. On the other hand, under public random monitoring the results are not so strong. However, in some circumstances, it is possible to enlarge the spectrum of monitoring intensities where payoffs above the static Nash can be sustained. Proposition 17 identifies the conditions for random monitoring efficiency improvements.

In the imperfect monitoring case, the analysis is restricted to SSE. It would be interesting to study the effects of random monitoring in asymmetric equilibria.

Finally, we stress the potential of applications that random monitoring presents in the context of dynamic incentives, mechanism design, or even new developments in dynamic game theory.

[32] The optimal decision rule in the deterministic and the random monitoring case cross twice or once, respectively, for improvements of the type of Part (i) or (ii) of Proposition 17.

2.7 References

1. Abreu, D., P. Milgrom and D. Pearce (1991). "Information and Timing in Repeated Partnerships." Econometrica, 59, 1713-1733.

2. Abreu, D., D. Pearce and E. Stacchetti (1986). "Optimal Cartel Equilibria with Imperfect Monitoring," Journal of Economic Theory, 39, 251-269.

3. Abreu, D., D. Pearce and E. Stacchetti (1990). "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." Econometrica, 58, 1041-1063.

4. Bernanke, B., and M. Gertler. (1989), "Agency Costs, Net Worth, and Business Fluctuations", American Economic Review, 79, 14-31.

5. Border, K. and J. Sobel. (1987). "Samurai Accountant: A Theory of Auditing and Plunder." Review of Economic Studies, vol.4, 525-540.

6. Faingold, E. (2006). "Building a Reputation under Frequent Decisions," mimeo.

7. Faingold, E., Y. Sannikov (2007). "Reputation Effects and Equilibrium Degeneracy in Continuous-Time Games," mimeo.

8. Friedman, J. W. (1971). "A Noncooperative Equilibrium for Supergames," Review of Economic Studies, 38, 1-12.

9. Fudenberg, D. and D. Levine (1994) "Efficiency and Observability with Long-Run and Short-Run Players," Journal of Economic Theory, 62(1), 103-135.

10. Fudenberg, D. and D. Levine (2007) "Continuous Time Models of Repeated Games with Imperfect Public Monitoring." Review of Economic Dynamics, 10(2), 173-192.

11. Fudenberg, D. and D. Levine (2009) "Repeated Games with Frequent Signals." Quarterly Journal of Economics, forthcoming.

12. Fudenberg, D., D. Levine and E. Maskin (1994). "The Folk Theorem with Imperfect Public Information." Econometrica, 62, 997-1040.

13. Fudenberg, D. and E. Maskin (1986). "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," Econometrica, 54, 533-554.

14. Fudenberg, D. and W. Olszewski (2009) "Repeated Games with Asynchronous Monitoring," mimeo.

15. Fudenberg, D. and J. Tirole (1991) Game Theory, MIT Press, Cambridge, MA.

16. Green, E. and R. Porter (1984). "Noncooperative Collusion under Imperfect Price Information." Econometrica, 52, 87-100.

17. Khalil, F. (1997). "Auditing without Commitment", RAND Journal of Economics, vol. 28, 629-640.

18. Mailath, G. and L. Samuelson (2006) Repeated Games and Reputations: Long-run Relationships. Oxford University Press, New York.

19. Monnet, C. and E. Quintin. (2005). "Optimal Contracts in a Dynamic Costly State Verification Model," Economic Theory, vol. 26: 867-885.

20. Mookerjee, D. and I. Png. (1989). "Optimal Auditing, Insurance and Redistribution," Quarterly Journal of Economics, vol. 104: 399-415.

21. Osório-Costa, A. M. (2008) "Frequent Monitoring in Repeated Games under Brownian Uncertainty," mimeo.

22. Porter, R. (1983). "Optimal Cartel Trigger Price Strategies." Journal of Economic Theory, 29, 313-338.

23. Radner, R., R. Myerson, and E. Maskin (1996) "An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria," Review of Economic Studies, 53, 59-69.

24. Sannikov, Y. (2007). "Games with Imperfectly Observable Actions in Continuous Time," Econometrica, 75, 1285–1329.

25. Sannikov, Y. and A. Skrzypacz (2007) "Impossibility of Collusion under Imperfect Monitoring with Flexible Production," American Economic Review, 97, 1794-1823.

26. Sannikov, Y. and A. Skrzypacz (2009) "The Role of Information in Repeated Games with Frequent Actions," Econometrica, forthcoming.

27. Stahl, II, D. O. (1991) "The Graph of Prisoners' Dilemma Supergame Payoffs as a Function of the Discount Factor," Games and Economic Behavior, 3, 368-384.

28. Townsend, R. (1979) "Optimal Contracts and Competitive Markets with Costly State Verification," Journal of Economic Theory, 21(2): 265-93.

2.8 Appendix - Proofs of the Lemmas and Propositions.

Proof of Proposition 11. Assuming all the conditions for pure-strategy SPE are satisfied; in the deterministic monitoring case, enforceability of the profile a, with respect to continuations in $V(\Delta, r)$, requires that

$$\delta^{\Delta} \geq \sup_{i\in N \text{ and } a_i' \in A_i} \frac{\pi_i(a_i', a_{-i}) - \pi_i(a)}{\pi_i(a_i', a_{-i}) - \pi_i(a) + w_i(a) - w_i(a_i', a_{-i})} \equiv \overline{\delta},$$

which is obtained from solving (27). Where $a_i' \in A_i$ is the most profitable deviation for player $i \in N$, along the equilibrium path. With $\pi_i(a_i', a_{-i}) \geq \pi_i(a)$ and $w_i(a) \geq w_i(a_i', a_{-i}) \geq \underline{v}_i^p$ where $\underline{v}_i^p \equiv \min_{a_{-i}\in A_{-i}} \max_{a_i \in A_i} \pi_i(a_i, a_{-i})$ is i's minmax payoff. Fix the continuation values $w(a)$ and $w(a_i', a_{-i})$ and assume that they are also in $\widetilde{V}(\Delta, r)$. Similarly, when monitoring is random, we solve (27) with δ^{Δ} substituted by $E_X(\delta^x)$, to obtain that $E_X(\delta^x) \geq \overline{\delta}$. Set $\delta^{\Delta} = \overline{\delta}$, then for the same equilibrium payoff v in $V(\Delta, r)$ and in $\widetilde{V}(\Delta, r)$, the expected discount factor under random monitoring is higher than the deterministic discount factor if $E_X(e^{-rx}) \geq e^{-rE_X(x)} = \delta^{\Delta}$, which is always true by Jensen's inequality when δ^x is convex in x, since $\Delta = E_X(x) > 0$. Then, we can increase r in $E_X(e^{-rx})$ to $\widetilde{r} > r$, such that $E_X\left(e^{-\widetilde{r}x}\right) = e^{-rE_X(x)} = \overline{\delta}$. Consequently, $V(\Delta, r) \subseteq \widetilde{V}(\Delta, r)$ and $w(.) \in \widetilde{V}(\Delta, r)$. ■

Proof of Corollary 13. When $\delta^x \in (0, 1)$ is strictly convex in x and $E_X(x) = \Delta$, by Jensen's inequality we have $E_X(\delta^x) > \delta^{\Delta}$. Then there might exist equilibria where

$E_X(\delta^x) > \overline{\delta} > \delta^{\Delta}$, for some common combination of r and Δ. The converse requires $\delta^{\Delta} > E_X(\delta^x)$, which cannot be true if the convex requirement of Proposition 11 holds. Then $V(\Delta, r) \subseteq \widetilde{V}(\Delta, r)$ for a same r and Δ. ■

Proof of Lemma 16. We will show the result for part (ii), part (i) follows the same lines. To find the expression that characterizes the best SSE payoff, we need to solve the following dynamic programing problem, which by symmetry is the same for both players:

$$\widetilde{v} = (1 - E_X(\delta^x))\,\pi + (1 - \widetilde{p}_{\delta}(\Delta))\,\widetilde{v} + \widetilde{p}_{\delta}(\Delta)\,\underline{\widetilde{v}}, \tag{40}$$

$$\widetilde{v} \geq (1 - E_X(\delta^x))\,\pi' + (1 - \widetilde{q}_{\delta}(\Delta))\,\widetilde{v} + \widetilde{q}_{\delta}(\Delta)\,\underline{\widetilde{v}}, \tag{41}$$

$$\underline{\widetilde{v}} = (1 - E_X(\delta^x))\,0 + E_X(\delta^x)\left[\widetilde{\alpha}\underline{\widetilde{v}} + (1 - \widetilde{\alpha})\,\widetilde{v}\right], \tag{42}$$

$$\widetilde{\alpha} \leq 1. \tag{43}$$

With $\widetilde{p}_{\delta}(\Delta)$ and $\widetilde{q}_{\delta}(\Delta)$ defined respectively in (30) and (31). An observation of a realized signal, inside the support Y^{-}, is interpreted as a signal of defective behavior. See the examples in the main text.

Expression (40) is the value of the relation when both players cooperate. Each player receives the expected normalized payoff associated with mutual cooperation, plus a discounted expectation over the expected values $\widetilde{v}$ and $\underline{\widetilde{v}}$, associated with the two types of signals that might be observed. The constraint (41) is an enforceability condition. The expected value of the game associated with mutual effort has to be at least as good as the expected value of the game associated with a potential unilateral deviation. By Definition 10, when satisfied, (40) and (41) enforce the profile (C, C).

Expression (42) is the value of the punishment phase, where $\widetilde{\alpha}$ is a number smaller than or equal to 1. See the technical discussion in the main text. We can look at $\widetilde{\alpha}$ as the probability of a correlated public signal which recommends that the relation remain in the punishment stage, but not exclusively. A value $\widetilde{\alpha} = 1$ means perpetual punishment, while $\widetilde{\alpha} = 0$ implies a single punishment period. When $\widetilde{\alpha} < 0$, a future compensation is required because of the excessive past punishment period. Providing condition (43) holds, we have $\underline{\widetilde{v}} \in [0, E_X(\delta^x)\,\widetilde{v}] \subset [0, \widetilde{v})$. Since (D, D) is a Nash equilibrium, punishment is trivially enforced.

Expression (42) is solved for $\underline{\widetilde{v}}$ to obtain

$$\underline{\widetilde{v}}(\Delta) = \frac{E_X(\delta^x)(1-\widetilde{\alpha})\widetilde{v}}{1-E_X(\delta^x)\widetilde{\alpha}}. \tag{44}$$

Plug $\underline{\widetilde{v}}(\Delta)$ into (40) and (41), with the latter holding with equality. We obtain (35), i.e.

$$\frac{1}{E_X(\delta^x)} - \frac{\widetilde{q}_\delta(\Delta)\pi - \widetilde{p}_\delta(\Delta)\pi'}{E_X(\delta^x)(\pi'-\pi)} = \widetilde{\alpha}(\Delta).$$

By (43) we must have $\widetilde{\alpha}(\Delta) \leq 1$ satisfying (35). Otherwise, no equilibria, other than the infinite repetition of the static Nash, can be sustained, i.e. $\widetilde{v}(\Delta) = 0$.

Replacing (44) and (35) into (40), we can solve for $\widetilde{v}$ to obtain the expression of the best SSE

$$\widetilde{v}(\Delta) = \pi - \frac{\widetilde{p}_\delta(\Delta)}{\widetilde{q}_\delta(\Delta) - \widetilde{p}_\delta(\Delta)}\left(\pi' - \pi\right).$$

Finally and for completeness, replacing $\widetilde{v}(\Delta)$ into (44), we obtain

$$\underline{\widetilde{v}}(\Delta) = E_X(\delta^x)\frac{\widetilde{q}_\delta(\Delta)\pi - \widetilde{p}_\delta(\Delta)\pi'}{\widetilde{q}_\delta(\Delta) - \widetilde{p}_\delta(\Delta)}\frac{1-\widetilde{\alpha}(\Delta)}{1-E_X(\delta^x)\widetilde{\alpha}(\Delta)},$$

which cannot be lower than 0, while $\widetilde{\alpha}(\Delta) \leq 1$ exists.

Part (i) is essentially the same with $E_X(\delta^x)$, $\widetilde{p}_\delta(\Delta)$ and $\widetilde{q}_\delta(\Delta)$ replaced by δ^Δ, $\delta^\Delta p(\Delta)$ and $\delta^\Delta q(\Delta)$, respectively. ■

Proof of Proposition 17. (i) When, for a given monitoring intensity $\Delta = E_X(x)$ and a common discount rate r, we have $\widetilde{\alpha}^*(\Delta) \leq 1$ and $\alpha^*(\Delta) \leq 1$, then both $\widetilde{v}^*(\Delta) \geq 0$ and $\overline{v}^*(\Delta) \geq 0$, i.e. both SSE payoffs attain values at least weakly above the static Nash. Then the profile (C, C) is enforceable with random and deterministic monitoring, and we need to compare the value of the best SSE payoff, i.e. $\widetilde{v}^*(\Delta) \geq \overline{v}^*(\Delta)$. The comparison between expressions (32) and (34), after some arrangements, gives condition (36).

(ii) When, for a given monitoring intensity Δ, and discount rate r, under optimal behavior, we have $\widetilde{\alpha}^*(\Delta) \leq 1$ and $\alpha^*(\Delta) > 1$, then $\widetilde{v}^*(\Delta) > \underline{\widetilde{v}}^*(\Delta) \geq 0$ and $\overline{v}^*(\Delta) = \underline{v}^*(\Delta) = 0$. That is, with deterministic monitoring in pure strategies, only the infinite repetition of the static Nash is enforceable. While random monitoring, allows for payoffs at least above the static Nash. Then $\widetilde{v}^*(\Delta) \geq 0 = \overline{v}^*(\Delta)$, i.e. random monitoring improves over deterministic monitoring. ■

Proof of Corollary 19. The expected discount factor effect always favors random monitoring, if, in addition, the informativeness of the public signals increases, then condition (36) necessarily holds. However, condition (36) might be satisfied even when (37) is not. The reason is that the expected discount factor effect might be stronger than the informational adverse effect in the public signals. ■

CHAPTER III

REPEATED INTERACTION AND THE REVELATION OF PLAYER'S TYPE: A PRINCIPAL-MONITOR-AGENT PROBLEM

3.1 Introduction

Inside firms, delegation is a common and natural practice as the organizational structure grows large. Delegation is the assignment of authority from one organizational level to a lower one. In the manufacturing industry, it is common to employ a monitor to supervise the activity of a set of blue collar workers. Still, the high rank manager remains accountable for the outcome of the monitor's and the agent actions which affect the firm's performance.

A distinct structure of delegation occurs in the auditing industry and banking/finance supervision, typically an auditor/supervisor is sent to the client office to look at his accounts and other documents that might provide useful evidence about the client's behavior. The payoff of the auditing company or the supervising authority can be measured in monetary or in reputational terms. In either case, the auditor/supervisor actions affect the value of these variables.

A monitor is an individual with personal characteristics which interfere with his professional performance. Returning to the manufacturing industry example, when a worker effort is observed with noise, two different monitors may disagree on whether a given output realization is a signal of high or low effort.[1]

The table below shows the actual effort choices made by a worker and the possible interpretations that a particular monitor can make when these choices are not perfectly

[1] There are some connections to subjective performance evaluations. See for example Baker, Gibbons and Murphy (1994), Bull (1987), MacLeod (2003) or MacLeod and Malcomson (1989). The main issue here is how a low signal of the agent effort is interpreted in terms of actual effort choice.

observable.

	Monitor Perception	
Worker Choice	High Effort	Low Effort
High Effort	*Correct*	*Type I Error*
Low Effort	*Type II Error*	*Correct*

For example, some monitors might be more permissive or tolerant with respect to their subordinates than others. Such is not necessarily a bad character trait. However, they are more likely to be the object of strategic behavior from the workers, because they tend to incur more often in type II errors. Strict or demanding individuals are also not necessarily better or worse in professional terms. They tend to incur more often in wrong type I judgements, but at the same time, they bring more discipline to the relation.

Typically, after a sequence of low performance observations, the monitor might decide to punish the particular worker in question. These punishments can take several forms; one is the possibility of firing the individual. In spite of this, to satisfy their natural tendency for low effort, workers may explore the monitor's personal characteristics and limitations, patterns of behavior and working methods. This information might be of strategic relevance, allowing them to revise their initial prior beliefs and readjust their behavior accordingly. Repeated interaction facilitates this potential corrosive learning, with negative impact on the firm's performance.

In the auditing/supervision industry there are monitoring standards that have to be followed. However, different supervisors may differ with respect to specific aspects and working methods. Bernard Madoff, known to have run the largest Ponzi scheme in world history, describes in the following way his experience with two different supervisors from the SEC at different moments in time in the following way:[2]

Madoff stated that Mr. X was "doing things that make no sense to us at all."... Mr. X "talked tough, but didn't look at anything".

[2] *The words in italics were taken from the description of an interview conducted by Inspector General H. David Rotz and Deputy Inspector General Noelle Frangipane with Bernard Madoff on June 17, 2009 about interaction between Madoff, his company and the U.S. Securities and Exchange Commission (SEC).*

Madoff ... recalled Mr. Y was the supervisor.... Mr. Y, "knew what he was looking at and that was it."

In Madoff's words it is clear that Mr. Y's working methods were different from the ones employed by Mr. X. More relevant for us is that Madoff, through observation of the both supervisors' actions, was able to rank them in terms of the likelihood of uncovering his Ponzi scheme. In fact, Madoff recognized that if the 2006 exam conducted by Mr. X had been conducted by Mr. Y, the Ponzi scheme would have been found. Translated to the setting of the present paper, Madoff would have more incentives to continue with his Ponzi scheme if he learns that the supervisor is of Mr. X type.

In general terms a fund manager that is rewarded with a performance fee would have more incentives to build a risky portfolio if he knew in advance what risk measures a particular supervisor would pay more or less attention to. Similarly, an unscrupulous manager would feel more tempted to enter into illicit activities of a given kind if he knew beforehand that the auditor would pay little or not attention to issues of that nature.

Blue collar workers and most managers are typically paid according to a flat compensation. Bonus or performance stimulus are possible in the case of good performance, but typically they do not share the losses. This fact limits, to a great extent, the provision of incentives. Monitoring is then the mechanism that disciplines these agents.

This paper studies a dynamic *principal-monitor-agent* relation where a strategic principal delegates the task of monitoring the effort choices of an agent to a third party. The latter we call the *monitor*. His actions are fully characterized by his type. Exogenous circumstances require the delegation of the monitoring activities, without which trade would not be possible. The agent is strategic and has a natural tendency to supply low effort.[3]

Through repeated interaction and the observation of the monitor's actions with respect to the effort signal, the agent may learns his type, which in our setting represents the flexibility of the monitor towards the observations of the realized output. The agent strategically

[3] The model history is particularly tailored to capture the firm versus blue collar workers type of relations. With appropriate changes in the text, the generalization for the auditing and financial supervision problems is immediate.

lowers his effort if he finds that the monitor is *tolerant*. We show that this revelation process damages the principal's payoff. When the principal strategic influence is restricted to deciding exclusively on whether or not to trade with the agent, we are able to characterize the worst case scenario.

In order to solve the principal's problem, we formalize the idea of replacement strategies, i.e. the principal replaces the monitor when she finds it convenient, paying a cost but disrupting any learning that the agent might have acquired.[4] When the replacement costs are null, she obtains the largest possible payoffs that can be achieved with replacement strategies. We are thus able to establish upper and lower bounds on the payoffs that both parties can achieve independently of the information structure, for different degrees of impatience.

In any realistic setting, replacement strategies cannot fully solve the principal's problem when contrasted with a reference measure. Surprisingly, this is true even if replacement costs are zero and high effort is always played in equilibrium. The reason is that, for the replacement mechanism to work, the parties cannot benefit from the potential revelation of a tolerant monitor - which in our setting is preferred in payoff terms when incentives are met, since he incurs less often in mistaken punishments.

Nonetheless, replacement strategies turn out to be useful to solve the principal's problem, reducing the losses associated with the agent learning and enlarging the spectrum of discounting, where equilibrium effort can be sustained. The success of these strategies depends crucially on the replacement costs and on the agent impatience.

We also characterize the sequential equilibrium under public and private monitoring, for varying replacement costs and impatience levels. When the noise signal of the agent effort is publicly observed the principal is able to make more precise replacement choices, because she knows exactly the moment in time at which the agent learns that the monitor type is tolerant. However, when the realized output is the agent's private monitoring, this piece of information is not available anymore; she has to infer the informational state of the relation.

[4] Holmström (1982) and Cripps, Mailath and Samuelson (2004) in different contexts show the existence of a similar revelation effect. They also mention the possibility of permanent replacement as a potential mechanism to solve the problem.

For that reason her replacement choices are always limited either by being premature, in the sense that the agent is still uninformed, incurring in an unnecessary cost, or by being late, in the sense that the monitor type is already revealed and the agent is providing low effort. Consequently and not surprisingly, replacement strategies are payoff inferior under private observation than under public monitoring.

This paper contributes to the theory of incentives by presenting replacement strategies as a mechanism to solve problems where an agent might acquire information with strategic value that penalizes the principal's payoffs. This paper provides recommendations on how a principal should optimally rotate a monitor in situations of this type. Such a solution is particularly relevant in situations where compensation is exogenously determined, limiting the provision of incentives to a great extent. In multiagent situations, the power of replacement strategies is amplified.

Discussion on the Main Assumptions - There is a set of persistent facts which this paper attempts to capture. They justify important specificities of the present paper and, in some sense, novel departures from the existing literature.

We introduce a distribution of monitor types, differing with respect to their flexibility towards the observations of the noisy effort measure. These individuals are not necessarily the ideal choice. The scarcity of "perfect profiles", the subjectivity and not necessarily well defined characteristics of an "ideal type" are not easy to identify, together with restrictions and biases in the recruiting process might lead to a selected candidate that is simply perceived as the best of a limited pool of screened individuals.[5]

Firms are aware of these limitations, but they also know the necessity of hiring these individuals for the regular functioning and expansion of their businesses.[6] These arguments rationalize the existence of the monitor in our model.

[5] The classical "secretary problem" and its multiple extensions provides sufficient intuition about the recruiter screening difficulties. See Ferguson (1989) and Freeman (1983).

[6] Leibenstein (1987) discusses a great number of inefficiencies that exist inside organizations.

We restrict the role of compensation as an incentive mechanism. The assumption parallels the vision of Alchian and Demsetz (1972) of monitoring as a way of providing the agent with incentives.[7] The monitoring role is motivated by the independence between compensation and performance. Job insecurity is the mechanism that disciplines the relation.

This environment comes as natural in many economies and industrial sectors where base salaries are determined by social norms and political or legal aspects.[8] It also captures rigidities observed in the labor markets. When unsatisfied with an employee's performance, the employer is more prone to fire him than to decrease his compensation.

We assume that independently of the performance, once employed, the worker receives the promised end of period compensation. This is a "no-slavery" condition that the principal must respect ex-post. Punishments are executed after the compensation is paid.

The strategic aspect of firms' behavior tends to be mainly directed towards their costumers and competitors. In relation with its employees, firms usually pay the agreed compensation and only fail to do so in special circumstances (e.g. financial difficulties, bankruptcy, etc.) that have little to do with incentives. A failure in commitment with an employee is not only seen by all the other employees but also spreads outside the firm's halls. Reputational considerations of this kind provide theoretical foundations for this assumption.

We abstain from discussing potential renegotiation from the initial agreement. This allows us to focus on the main issue of this paper without extra complexities.

Related Literature - The revelation of a player's type is not a new issue. However, it has never been applied to a principal-monitor-agent relation with the kind of structure as in this paper.

In a setting where incentives are driven by career concerns, Holmström (1982a) shows that an individual's ability is revealed over time through observation of his performance.

[7] See also Shapiro and Stiglitz (1984) where they link wages, unemployment, monitoring and efficiency. The worker incentives for high effort are also induced by the fear of being fired.

[8] Clearly, this is not an universal fact, it is easy to find exceptions, in particular in white-collar jobs where the employee has attributes/skills, recognized by both parties, that endow him with bargaining power or in tasks where performance and compensation are linked in a very sensitive way. Even in these cases, we should also expect a market reference measure to base the negotiations. See Kennan and Wilson (1989) for a survey on strategic bargaining models between unions and firms and for empirical studies.

On the limit, this result holds no matter how noisy the performance measure is and no matter how much the agent tries to bias the principal beliefs.[9]

Cripps, Mailath and Samuelson (2004) recently showed that if monitoring is imperfect, the type of player that is building a reputation is revealed in the long-run. Either the opponent players will become convinced that the reputation builder strategy will not change or they will come to learn his type. The former case cannot be an equilibrium; otherwise the reputation builder would take advantage of that fact to deviate from the potentially costly reputational action. In equilibrium, the revelation of the true reputation builder type occurs almost surely.[10]

One common characteristic that is reported in these papers is that the revelation of the player's type has negative payoff consequences for the revealed player. In order to deal with the problem, a solution that has been proposed by these authors is the permanent replacement of the player whose type is initially unknown, disrupting any potential for learning.[11] Such a costless solution does not fit and is hard to motivate in many economic problems. The present paper extends these ideas for different informational structures, in a simple setting where the replacement possibility is costly.

The literature in three party relations, principal-supervisor-agent, is in particular concerned with the resolution of the potential breakout of collusive arrangements between the agent and the supervisor. Tirole (1986) points out that problems related with the monitor's conflicting interests might arise. See also Laffont and Tirole (1991) and Kofman and Lawarree (1993) for further developments and extensions on this literature. In this paper, we are not concerned with delegation effects of this kind. Instead, we focus our attention

[9] Fudenberg and Tirole (1986) present an explicit theory of predation where an incumbent firm attempts to bias the learning process of an entrant firm about the market profitability. See also Mirman, Samuelson and Urbano (1993).

[10] See Cripps, Mailath and Samuelson (2007) where they establish analogous results for the case where the uninformed player is long-lived. See also Cripps, Ely, Mailath and Samuelson (2008). Revelation issues due to repeated interaction are also common in dynamic games with incomplete information; Renault (2009) provides a survey on the topic.

[11] Cole, Dow and English (1995) and Phelan (2006) model situations where the government type in power is not permanent. These situations are easy to motivate. See also Mailath and Samuelson (2001), where they study a problem where a competent firm might become inept, this mechanism keeps the competent firm with incentives and is an equilibrium.

on the negative effects associated with the revelation of strategic relevant information.

This paper is also related to the theory of self-enforced contracts where the provision of incentives is guaranteed by the sensitivity of the continuation value of the infinitely repeated game to changes in the player's actions. Future rewards and punishments provide the incentives for present behaviour.[12] Some important contributions to the large and growing literature in relational contracts are Klein and Leffler (1981), Bull (1987), and MacLeod and Malcomson (1989), all these papers assume perfect monitoring; while Shapiro and Stiglitz (1984), Baker, Gibbons and Murphy (2002), Levin (2003) and Fuchs (2007), assume imperfect monitoring. The present paper is in the spirit of the former set of contributions since we assume a different spectrum of monitoring imperfections.

Levin (2003) shows, that an optimal relational contract has an equivalent stationary representation that achieves the same payoffs. Because of the uncertainty about the moment in time when the monitor type is revealed (if it is revealed) and due to the replacement possibility, multiple alternations between informational sates are possible. For that reason, Levin's result does not generalize to our model; we have to solve the dynamic problem with dynamic constraints.

The paper is organized as follows. Section 3.2 presents the general model structure and discusses the learning process. Section 3.3 discusses the principal worst case scenario in detail and defines the reference measure. Section 3.4 introduces the possibility of the monitor being replaced and considers the costless replacement solution. Section 3.5 characterizes the equilibrium replacement strategies under public monitoring. Similarly, Section 3.6 characterizes the case where the agent privately observes the noisy effort signals. Section 3.7 concludes. Proofs for the results in this paper can be found in Appendix 3.9.

3.2 The Model

We model a dynamic *principal-monitor-agent* relation that incorporates both elements of adverse selection and moral hazard. The former is motivated by the existence of a monitor,

[12] The present paper mixes concepts both from the theory of repeated games and the incentives theory. See Fudenberg and Tirole (1991) and Mailath and Samuelson (2006) for complete surveys of the former theory, and see for example Bolton and Dewatripont (2005) and Salanié (2005) for surveys of the latter theory.

whose type is initially unknown to the agent and to the principal. We assume that the monitor has no incentives to choose an action that is against his type. Consequently, the monitor's actions are revealing. The problem can then be treated as one of dynamic moral hazard where nature first selects the monitor type.

The Principal and the Agent - Both parties are risk neutral. The timing of the relation is the following: at the beginning of each period $t = 0, 1, 2...$, the principal offers the fixed market compensation $w_t \in \mathbb{R}_+$ in exchange for the agent effort. The latter has the option to accept or refuse it. Once accepted, the agent privately chooses an effort level $e_t \in \left\{e^L, e^H\right\}$, where e^H denotes high effort and e^L denotes low effort, and $e^H > e^L \geq 0$. The effort level e^H is exogenous and it is the highest effort that the agent can physically supply per period. The low effort level e^L will be endogenously determined below.

At the end of each period, before players choose their actions, the realized output $y_t \in \mathbb{R}$ is observed by the monitor and the agent.[13] The principal might observe it or not, depending on the information structure considered. The realized output is a noisy measure of the agent's effort, i.e. $y_t = be_t + \varepsilon_t$ where ε_t is stochastic with $E(\varepsilon_t) = 0$ and $b \in \mathbb{R}_+$ is a productivity parameter. High (low) effort choices make higher (lower) outputs more likely.

The principal's ex-post payoff is the realized output subtracted from the wage paid to the agent. The principal's ex-ante expected payoff is then

$$\pi_{pt}^{j} = be_t - w_t, \tag{45}$$

with $j = L$ when $e_t = e^L$ and $j = H$ when $e_t = e^H$.

Independently of the observed output, the principal has necessarily to pay the wage w_t corresponding to that period. We look at the principal as having one period commitment. She has, however, the option of not starting the relation with the agent or terminate it, if she finds convenient.

In exchange for the compensation paid at the end of the period, the agent suffers a disutility from effort equal to ce_t, where $c \in \mathbb{R}_+$ is the marginal cost of effort. The agent

[13] There is no loss in generality when allowing for potentially negative output values.

payoff is given by

$$\pi_{at}^{j} = w_t - ce_t, \tag{46}$$

with $j = L$ when $e_t = e^L$ and $j = H$ otherwise. Assume $b > c$, to assure an expected positive surplus.

Notice that there is no uncertainty about the agent's per period payoff. The principal holds all the idiosyncratic production risk.

If there is no trade both parties obtain their respective outside options denoted as $\underline{v}_p$ and $\underline{v}_a$. There is a wide range of interpretations that can be given to the values $\underline{v}_a$ and $\underline{v}_p$, we let these values to be exogenous.[14]

In the stage game, low effort is preferred by the agent but not by the principal. In order for the agent to always have an incentive to participate, we assume $\pi_a^L > \pi_a^H > \underline{v}_a$. The principal stage game payoff follows the relation $\pi_p^H > \underline{v}_p \geq \pi_p^L$. We assume that the last inequality binds, i.e. $\underline{v}_p = \pi_p^L$. This assumption simplifies the problem later. However, it has implications on the stage game.

	e^H	e^L	$\varnothing$
E	π_p^H, π_a^H	$\underline{v}_p, \pi_a^L$	$\underline{v}_p, \underline{v}_a$
NE	$\underline{v}_p, -ce^H$	$\underline{v}_p, -ce^L$	$\underline{v}_p, \underline{v}_a$

The principal is the row player, who has the option of employing E or not employing the agent NE. The agent is the column player, together with the effort choices he has the option of not trading with the principal, denoted as $\varnothing$, guaranteeing the outside option $\underline{v}_a$. With the assumption $\underline{v}_p = \pi_p^L$, instead of a single pure strategies Nash equilibrium $(NE, \varnothing)$ we have two, i.e. $\left(E, e^L\right)$ and $(NE, \varnothing)$. None of these equilibria guarantee the principal more than the outside option. The equilibrium $(NE, \varnothing)$ is not interesting for either party but the equilibrium $\left(E, e^L\right)$ is preferred by the agent. The assumption is without loss of generality because, in the repeated relation, we want to sustain the repetition of the outcome

[14] These values can be easily endogenized at the cost of a more complex problem. For example; $\underline{v}_p$ could represent the value of a new relation with another agent, subtracted from the associated searching costs. A similar interpretation can be done for $\underline{v}_a$.

associated with the equilibrium $\left(E, e^H\right)$.[15]

The Monitor - The monitor's task is to supervise the agent effort. He has autonomy to punish the agent when he observes a low output realization and subsequently reports this event to the principal.

The employed monitor can be of two types $\theta \in \left\{\theta^T, \theta^S\right\}$, with $\theta \in \mathbb{R}$ and $be_t^H > be_t^L \geq \theta^S > \theta^T$.[16] The type θ^S occurs with probability $\beta \in (0,1)$ and the type θ^T with the remaining probability. Here θ^S denotes a "strict" monitor, i.e. an individual that is more likely to perceive a low output realization as a signal of low effort. The strict type considers a low output realization every signal $y_t \leq \theta^S$, i.e. $Y^S \equiv \left\{y_t : y_t \leq \theta^S\right\}$, and a high output realization otherwise. A less strict monitor, call it "tolerant", is denoted by θ^T. This type attempts to capture a more flexible individual towards the output observations, so low output realizations are less likely to be interpreted as signalling low effort. For a tolerant monitor any output in the set $Y^T \equiv \left\{y_t : y_t \leq \theta^T\right\}$ is a signal of low effort.

Consequently, when the agent provides high effort, low output is observed with different probabilities depending on whether the monitor is "tolerant" or "strict",[17] i.e.

$$\Pr\left(y_t \in Y^T | e_t = e^H\right) \equiv p^T \textit{ and } \Pr\left(y_t \in Y^S | e_t = e^H\right) \equiv p^S,$$

respectively. High output is interpreted with the remaining probabilities.

Similarly, when the agent chooses low effort, the output is low with probabilities

$$\Pr\left(y_t \in Y^T | e_t = e^L\right) \equiv q^T \textit{ and } \Pr\left(y_t \in Y^S | e_t = e^L\right) \equiv q^S,$$

depending on whether the monitor is "tolerant" or "strict", respectively.

To shorten notation, given the prior beliefs about the persistence of each type, we define

[15] Notice that the lower bound on effort supplied by the agent becomes endogenously determined, i.e. $e^L = \left(w + \underline{v}_p\right)/b$.

[16] The choice of two monitor types is made for simplicity. The model is robust to the introduction of more types and a continuum of efforts choices. The addition of more types simply increases the complexity of the problem.

[17] Without loss of generality we can consider another interpretation for θ, such as, for example, the state of nature. Clearly the story of the problem would have to be adjusted accordingly. Another interpretation is to look at θ as a monitoring technology, rather than an individual. The latter interpretation was suggested by Jacques Crémer. Nonetheless, we prefer to look at the monitor as a human.

the "expected" type $\theta^E \equiv \beta\theta^S + (1-\beta)\,\theta^T$.[18] In this case, the "expected" probability of observing low output when the agent is providing high and low effort are respectively

$$\beta p^S + (1-\beta)\,p^T \equiv p^E \textit{ and } \beta q^S + (1-\beta)\,q^T \equiv q^E. \tag{47}$$

Since $\theta^S > \theta^T$ we have $Y^T \subset Y^S$, then for a same effort choice, a low output interpretation is more likely when the monitor type is "strict",

$$p^S > p^E > p^T > 0 \;\; \textit{and } q^S > q^E > q^T > 0. \tag{48}$$

Within the same type, high effort has associated a lower punishment probability

$$q^T \geq p^T, q^S \geq p^S \textit{ and } q^E \geq p^E. \tag{49}$$

Putting together the inequalities (48) and (49) we obtain, without specifying an order, that q^T and p^S must lie in the interval $\left(p^T, q^S\right)$.

We ignore any payment made to the monitor by the principal and we do not specify the monitor's payoff functions. We assume that a "tolerant" monitor has no incentives to misbehave pretending to be of a "strict" type and vice versa. The value of θ not only denotes the type of monitor but also determines his behavior. The presence of the monitor is crucial, otherwise no trade would be possible and both parties would get their outside options. We see the monitoring task as more complex, specialized and with more responsibilities than simply observing output realizations. This assumption justifies the presence of the monitor even when the principal observes the realized output.

The Revelation Probabilities - The output observation carries a signal concerning the effort supplied by the agent. The monitor's reaction to the signal conveys information about his type to the agent. It is then natural to expect that a strategic agent would take advantage of this aggregated information. This is the intuition behind the revelation process.

[18] Outside the expected utility hypothesis, a more conservative approach would require the initial effort choice to be based on θ^S rather than on θ^E. There is no ambiguity about the likelihood of both types.

Definition 20 *Conditional on no punishment, in the event* $\left\{\theta^T < y_t \leq \theta^S\right\}$ *we say that a revealing signal has occurred.*

In the two types setting that we build in this paper, revealing information occurs only when the true type is θ^T. In this case the agent can update his believes and consequently revises his effort.

For simplicity, the event that the agent learns that the monitor is θ^S coincides with a punishment decision. Consequently, learning that the monitor is strict is irrelevant. Such is a consequence of punishment exclusively based on firing the agent.[19]

With two types, after observing a revealing signal and updating his beliefs there is nothing else for the agent to learn, so the type of monitor is fully revealed.

Remark. *If we added more types, conditional on no punishment, the agent would be learning more information about the monitor type with positive probability in every period. With punishment schemes based on the termination of the relation, the true value of* θ *could never be perfectly learned; the punishment disrupts this process.*[20] *However, the agent need not know exactly the type of monitor in order to discover profitable deviations.* ■

In every period of the relation with probability

$$\Pr\left(\theta = \theta^T\right)\Pr\left(y_t \in \underline{Y}^T \cap \underline{Y}^S | e_t = e^H\right) \equiv (1-\beta)\, r$$

the agent observes a revealing signal that excludes the type θ^S with probability one (in the spirit of Bayesian updating).[21] The larger the value of $(1-\beta)\, r$, the smaller the expected number of periods needed for a revealing signal to occur. In case of low effort e^L, the agent

[19] For such a case to be interesting, we would have to define an extra effort level, i.e. after the agent learns that the monitor is strict, he would adjust to a higher effort level. That would require the definition of more probabilities. Additionally we would have to consider other punishment schemes, e.g. review or forgiving strategies. See Footnote 20.

[20] *If the punishment allows for forgiveness and there is unbound recall, the learning process will not be disrupted. It will continue after the punishment has been completed. In this case,* θ *might become common knowledge after a sufficiently large number of periods. In this case, the agent can also learn the value of* θ *from below, by updating the lower bound on the distribution of monitor types. (A possibility that we are not considering here) This would accelerate the convergence to common knowledge. See Cripps, Ely, Mailath and Samuelson (2008). Such a setting requires different strategic considerations.*

[21] In general, this is different from saying that there is a signal clearly revealing the type of monitor. The former case does not necessarily lead to common knowledge, while the latter does. Here, with only two types, by exclusion both situations are equivalent. See the remark after Definition 20.

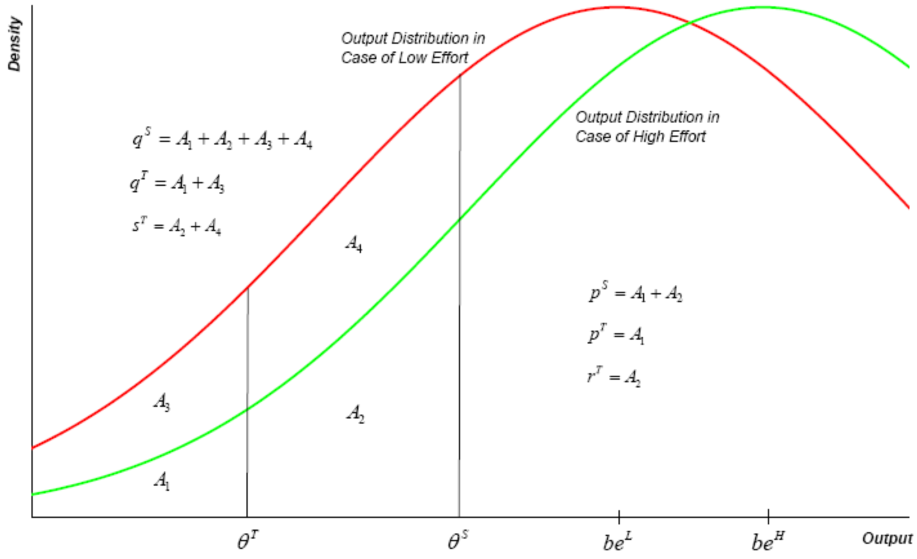

Figure 9: The model - Relation between revelation and punishment probabilities.

might observe a revealing signal with probability $(1-\beta)\,s$. We assume that $s \geq r$, this is the case for most distributions of interest. Consequently, when the agent provides low effort he accelerates the potential revelation process, but at the same time he is also more likely to be punished.

Since $be^L \geq \theta^S$, we can relate the punishment and the unconditional learning probabilities, i.e. $p^S = p^T + r$ and $q^S = q^T + s$. We can also write $p^E = p^T + \beta r$ and $q^E = q^T + \beta s$. Without knowing the relation between $q^S - q^T$ and r, we cannot rank q^T and p^S. Figure 9 provides an illustration.

Actions, Histories and Strategies - The Nash equilibria of the stage game gives low payoffs for the principal. Through the provision of intertemporal incentives, we can achieve payoffs above the set of stage game Nash equilibria.

For convenience, we consider strategies where a low output observation triggers the immediate termination of the relation. Both parts are then left with their outside options.[22] The punishment scheme interferes with the learning process, see Footnote 20. The important

[22] In the real world principal-agent relations punishments are usually not so severe. Typically, a firing decision is only made after a sequence or a certain number of low output interpretations. Radner (1985), Rubinstein (1979) and Rubinstein and Yaari (1983) study strategies with a similar structure.

aspect is that these punishments are mutual and occur with positive probability along the equilibrium path, causing destruction of value, and consequently bounding the equilibrium payoffs away from full efficiency. Even the reference value is not fully efficient. Instead of searching for a mechanism that implements an optimal effort level without surplus losses, we acknowledge the existence of inefficiencies. This assumption is without loss in generality and is motivated by tractability issues. Moreover, we assume that punishments based on a "money-back guarantee" cannot be contracted.[23]

The potential revelation of the monitor type creates an extra layer of inefficiency and damages the principal's interests. Our goal is to search for solutions that attempt to eliminate or minimize this latter effect on the principal's payoffs.[24]

Both the principal and the agent discount the future according to some $\delta \in (0, 1)$.

Once the principal and the agent decide to participate, the monitor is hired and the parties take their actions. At any given moment in time t; an agent's action is an effort choice $e_t \in \left\{e^L, e^H\right\}$. A principal's action is a replacement choice $r_t \in \{0, 1\}$, where $r_t = 0$ means not replacing the monitor and $r_t = 1$ otherwise. A monitor's action $m_t \in \{0, 1\}$ is a choice between punishing or not, respectively $m_t = 1$ and $m_t = 0$.

Given the actions and the observed output up to time t, a history of play is built. Depending on the information structure considered, different private histories players will accumulate. The history of output realization at a given time t is $h_y^t \equiv \{y_0, y_1, ..., y_{t-1}\}$. The monitor history of actions is $h_m^t \equiv \{m_0, m_1, ..., m_{t-1}\}$ and is public observed in any setting, for that reason we are able to keep the recursive structure. The principal and the agent condition their actions on this public observed history, even holding private and

[23] We assume that the only way to provide incentives is through mutual punishment. Alternatively, given the one-sided moral hazard structure of our problem, in theoretical terms we could consider other ways of providing incentives that are less costly for the parts, i.e. by transfers of value between the parties involved in the relation. Even with noisy signals about players' actions Fudenberg, Levine and Maskin (1994) (see also Fudenberg and Levine (1994) and Sannikov (2007)) have shown that, in problems of this type and with arbitrarily patient players, we can obtain any feasible and individual rational payoffs that are fully efficient. However, for our particular problem, transfers of value from employees to employers as a punishment scheme are not usually observed in reality.

[24] In a bilateral risk neutral setting, when the agent owns or receives all the surplus from his work, it is possible to obtain full efficient payoffs. The agent has no incentive to deviate from high effort. See for example Stiglitz (1974).

different pieces of information.[25] This way we can apply the dynamic programming methods developed in Abreu, Pearce and Stacchetti (1986, 1990). The principal and the agent private histories are $h_p^t \equiv \{r_0, r_1, ..., r_{t-1}\}$ and $h_a^t \equiv \{e_0, e_1, ..., e_{t-1}\}$ respectively.

Throughout the paper we focus on two information structures. In either case both the agent and the monitor observe the realized output. When the principal also observes the realized output, we say that *monitoring is public*; then the public history is $h_y^t \cup h_m^t$, the principal's private history is $h_y^t \cup h_m^t \cup h_p^t$ and the agent's private history is $h_y^t \cup h_m^t \cup h_a^t$. When the principal does not observe the realized output we say that the signals are the *agent's private monitoring*: then the public history is h_m^t, the principal's and the agent's private histories are respectively $h_m^t \cup h_p^t$ and $h_y^t \cup h_m^t \cup h_a^t$.

A pure strategy for player $i \in \{a, p\}$ is a mapping from the set of $i's$ private histories into the set of $i's$ pure actions. When the distribution of signals generated by the effort choices has full support, a *perfect public equilibrium* is a *sequential equilibrium.* In order to be general enough to deal with private monitoring structures, we work with the latter concept.

Notation - Players select the same action for every period of the repeated game until the monitor type is revealed or the monitor is replaced, in which case a new action is selected; for that reason we remove the time index t. Instead, we distinguish between the uninformed state, denoted with the superscript 0, and the informed state, denoted with the superscript 1. The punishment event is an absorbing state.

3.3 The Reference Value and the Revelation of the Monitor Type

In this Section we restrict the principal's strategic role other than deciding on whether or not to hire the agent and stoping the relation if she finds it convenient. This strategic structure is particularly interesting since it highlights the destructive effect that the revelation of the

[25] Kandori (2002) points out the difficulties that arise when dealing with private monitoring when a recursive structure is absent. Compte (1998) and Kandori and Matsushima (1998) present the first folk theorem for private monitoring with communication. See also Gossner and Tomala (2009) and Mailath and Samuelson (2006) for surveys on the subject.

monitor type has on the principal's payoff. This is the principal's worst case scenario.

The agent incentives are provided by the fixed market compensation and the uncertainty about the monitor type. The latter incentives might disappear as the game unfolds. With positive probability the agent might learn that the monitor type is tolerant and adjust his effort accordingly.

Given the strategic restrictions on the principal's behavior, whether she observes the realized output or not becomes irrelevant.

The Repeated Relation Payoffs - For $i \in \{a, p\}$, denote π_i^0 as the stage game payoff in the uninformed state 0, i.e. before the monitor type has been revealed. This value depends on the effort choices made by the agent, i.e. $\pi_i^0 \in \{\pi_i^L, \pi_i^H\}$. Similarly, in the informed state 1, i.e. when the monitor type has been found to be tolerant, the stage game payoffs are $\pi_i^1 \in \{\pi_i^L, \pi_i^H\}$. Let $v_i^{0,1}$ denote the normalized expected value, for $i \in \{a, p\}$, where the first superscript refers to the effort choice made in state 0 and the second superscript refers to the effort choice made in state 1.

Lemma 21 *Suppose the agent and the monitor observe the realized output. The infinitely repeated normalized expected payoff when the agent chooses $\{e^0, e^1\} = \{e^H, e^L\}$ is*

$$v_i^{H,L} = \frac{(1-\delta)\,\pi_i^H + \delta p^E \underline{v}_i}{D} + \delta\,(1-\beta)\,\frac{r}{D}\,\frac{(1-\delta)\,\pi_i^L + \delta q^T \underline{v}_i}{1-\delta\,(1-q^T)}. \tag{50}$$

for $i \in \{a, p\}$, where $D \equiv 1 - \delta\left(1 - p^E - (1-\beta)\,r\right)$.

Expression (50) has two components, the first ratio on the right-hand side is $i's$ ex-ante expected value of the uninformed state, when the agent provides high effort. The second part on the right-hand side is $i's$ ex-ante unconditional expected value of the informed state, when the agent provides low effort.

From expression (50) we can obtain the payoff associated with low effort and high effort in any informational state, $v_i^{L,L}$ and $v_i^{H,H}$ respectively. In the former case, replace respectively p^E, r and π_i^H for q^E, s and π_i^L. (Such an expression, for the case where $i = a$, can be found in the Proof of Proposition 23, expression (68).) The latter case is obtained by replacing respectively q^T and π_i^L for p^T and π_i^H in (50). (This is expression (56) of Definition 24, below.)

The effort choices $\{e^0, e^1\} = \{e^L, e^H\}$ are never optimal. In other words, low effort and an associated high probability of punishment while uninformed, is not compatible with a later high effort when the agent learns that the monitor is tolerant. In our setting, it can be shown that $v_a^{L,H}$ is always dominated by the payoffs of some other strategy for all $\delta \in (0, 1)$.

The Agent Effort Incentives - As discussed in Section 3.2, the principal pays the promised end of period compensation independently of the observed performance. The principal intertemporal incentives are then satisfied by assumption. The same does not happen with the agent who has a natural tendency to supply low effort. When the agent is sufficiently impatient the market compensation may not be enough to sustain high effort in both informational states. He provides high effort while uninformed and low effort once informed. It might even be the case, for larger impatience levels, that the agent finds it is in his best interest to provide low effort in any state.

We say that the agent effort choice $\{e^0, e^1\}$ is self-enforceable[26] when

$$\{e^0, e^1\} = \underset{e^0 \in \{e^L, e^H\},\ e^1 \in \{e^L, e^H\}}{\arg\sup} v_a^{0,1}. \tag{51}$$

The effort choices $\{e^0, e^1\}$ can be part of a non trivial equilibrium if, in addition, they guarantee that the principal has incentives to hire the agent. Before that, the following remark is in order, respecting to the methods employed to solve the dynamic problem that we are modelling.

Remark 22 *With two potential informational states, we can decompose the repeated relation into two relevant subgames. The whole game that starts in state* 0 *and extends for an unknown number of periods until the monitor interprets a low output realization, suggesting low effort. In this case the relation terminates. The state* 1 *subgame is initiated if the monitor type is revealed to be* θ^T*.*

[26] Throughout the paper we say that the agent has incentives to choose $\{e^0, e^1\}$ (or the agent incentives are satisfied). In the theory of incentives, expression 51 is an incentive compatible constraint.

When the informational state moves to state 1*, the agent might change his behavior. To solve the game, we first find the best strategy for the agent in any of the potentially infinitely repeated state* 1 *subgames and then search for the best strategy for the game that starts in the uninformed state* 0*. The approach is similar to backward induction; the difference is the timing uncertainty associated with the beginning of the informed state subgame.*

The Principal Participation - We require the ex-ante condition $v_p^{0,1} > \underline{v}_p$ as necessary for the principal to have interest in trading with the agent, i.e. the principal will only participate if she can do strictly better than in the stage game. However, this does not guarantee the principal's participation in the informed state 1 subgame. We denote the value of this subgame as $v_p^{\cdot,1}$. In the worst case scenario, i.e. the agent provides low effort in the informed state, we want

$$v_p^{\cdot,L} = \frac{(1-\delta)\,\pi_p^L + \delta q^T \underline{v}_p}{1-\delta\,(1-q^T)} \geq \underline{v}_p. \tag{52}$$

Otherwise, the principal would prefer to terminate the relation once the agent is informed.[27] For the principal to have incentives to stay active, the agent cannot provide effort below a certain value. For that reason, in Section 3.2, we assumed that $\underline{v}_p = \pi_p^L$. Intuitively once informed, if the agent decides to supply low effort, there is a minimum effort level that keeps the principal interested in participating.

With this restriction on the agent's effort, the principal's participation condition $v_p^{0,1} > \underline{v}_p$ is sufficient. When $v_p^{0,1} = \underline{v}_p$, the principal will not trade with the agent. This is the case when the agent's dominant strategy is to provide low effort in both informational states, i.e. $v_p^{L,L} = \underline{v}_p$. The principal does not contract with an agent who has no incentives to provide high effort at least during the uninformed state.[28]

The Agent Participation is always guaranteed by the assumption $\pi_a^L > \pi_a^H > \underline{v}_a$. In any scenario the agent guarantees at least $\underline{v}_a$.

[27] Termination of the relation after an informative signal can be part of an equilibrium. Here, we want to rule out such a possibility.

[28] We thus disallow strategic behavior from the low effort agents, so as to provide high effort in just one period of the relation, which would push the principal's payoff above $\underline{v}_p$ in expected terms. The principal participation decision is based on the agent's incentives, which the principal knows because δ and other relevant parameters are common knowledge. We want to improve over the stage game payoff $\underline{v}_p$.

Putting together the principal's and the agent's participation conditions, we obtain the constraints on the agent binary effort set

$$\frac{w+\underline{v}_p}{b} = e^L < e^H < \frac{w-\underline{v}_a}{c}, \tag{53}$$

with e^L endogenously determined and e^H exogenous but constrained. The former value is higher when the principal has a high outside option, the market compensation is high or the productivity of the labor is low. Also intuitive is the upper bound on high effort, which depends positively on the market wage but negatively on the agent's outside option and the cost of effort.

Equilibrium Behavior - Given that high effort is supplied in the uninformed state, let δ^+ denote the discount threshold above which the agent prefers to supply high effort once informed that the monitor type is θ^T. Given the principal's strategic limitation, δ^- denotes the discount threshold below which the agent prefers to provide low effort in any informational state.

Proposition 23 *When $\delta \in \left[\delta^-, \delta^+\right)$, the agent chooses $\left\{e^0, e^1\right\} = \left\{e^H, e^L\right\}$. The interval $\left[\delta^-, \delta^+\right) \subseteq [0,1)$ and is nonempty if*

$$q^T\left(\pi_a^H - \underline{v}_a\right) > p^T\left(\pi_a^L - \underline{v}_a\right) \tag{54}$$

and

$$s\left(\pi_a^H - \underline{v}_a\right) > r\left(\pi_a^L - \underline{v}_a\right). \tag{55}$$

Independently of the informational state, when $\delta \in \left[0, \delta^-\right)$ the agent provides low effort, while if $\delta \in \left[\delta^+, 1\right)$ the agent provides high effort.

The agent provides low effort in the informed state if $\delta < \delta^+$. Given this behavior, it is a dominant strategy to supply high effort in the uninformed state if $\delta \geq \delta^-$. Condition (54) and (55) establish that the intersection of these intervals is nonempty, while condition (54) alone guarantees that $\delta^+ < 1$.

Since $\pi_a^L > \pi_a^H > \underline{v}_a$, condition (54) states that it is not enough for q^T to be larger than p^T, it has to be sufficiently large. If that is the case, for a sufficiently patient agent, the market compensation is sufficient to sustain high effort.

Similarly, inequality (55) requires that the likelihood of an output observation inside the informative region to be much greater in the case of low effort. This condition is easier to interpret when added to condition (54). In this case we have $q^S\left(\pi_a^H-\underline{v}_a\right)>p^S\left(\pi_a^L-\underline{v}_a\right)$. Low effort accelerates the potential revelation of the monitor type, but at the cost of a higher likelihood of punishment. In other words, since a tolerant type occurs only with probability $(1-\beta)\in(0,1)$, the acceleration of the revelation process turns into an acceleration of the punishment event, independently of the true monitor type. This explains why the agent chooses $\left\{e^0,e^1\right\}=\left\{e^H,e^L\right\}$ for $\delta\in\left[\delta^-,\delta^+\right)$, but also why $\left\{e^0,e^1\right\}=\left\{e^L,e^H\right\}$ is always a dominated strategy.

Proposition 23 identifies three distinct potential behaviors depending on how the agent discounts the future. When $\delta\in\left[\delta^+,1\right)$, the agent provides high effort, ignoring any signal regarding the monitor type. The agent prefers to suffer the extra disutility $c\left(e^H-e^L\right)$ imposed by high effort in every future period of the repeated game rather than to provide low effort and increasing the probability of punishment from p^T to q^T. The reference payoff $v_a^{H,H}$ is dominant and the principal obtains the payoff $v_p^{H,H}>\underline{v}_p$. In the opposite direction, if $\delta\in\left[0,\delta^-\right)$, the agent incentives are not satisfied and the principal prefers not to participate since $v_p^{L,L}=\underline{v}_p$.

The interesting possibility occurs for "intermediate" impatience levels, i.e. $\delta\in\left[\delta^-,\delta^+\right)$. The agent prefers to provide high effort while uninformed, reducing the punishment likelihood, waiting to find if the monitor type is tolerant, in which case he deviates to low effort and obtains larger expected gains. There is a prize for the agent if the informed state occurs. In this case the principal obtains $v_p^{H,L}>\underline{v}_p$.

The Reference Payoffs - Since there is the potential of mistaken punishments on the equilibrium path, there is no way to achieve full efficient payoffs. To perform payoff comparisons we define the reference payoffs, i.e. the payoffs that attain the highest possible aggregate surplus for a given δ.

Definition 24 *The reference payoff is the value that the principal and the agent would* *obtain if the agent provides high effort independently of the informational state, i.e.*

$$v_i^{H,H} = \frac{(1-\delta)\,\pi_i^H + \delta p^E \underline{v}_i}{D} + \delta\,(1-\beta)\,\frac{r}{D}\,\frac{(1-\delta)\,\pi_i^H + \delta p^T \underline{v}_i}{1-\delta\,(1-p^T)}. \tag{56}$$

for $i \in \{a, p\}$ *.*

The reference payoffs require high effort always to be chosen by the agent, Which is why the outcome is less inefficient. Instead of attempting to eliminate this inefficiency we acknowledge it,[29] our goal is to minimize the principal's payoff losses due to the revelation of the monitor type.

It is important to stress, that $v_i^{H,H}$ is an equilibrium payoff only when $\delta \in \left[\delta^+, 1\right)$. For that reason, for any $\delta < \delta^+$, the value $v_i^{H,H}$ only plays the role of a reference measure. Nonetheless, the sum $v_a^{H,H} + v_p^{H,H} \equiv W$ achieves the largest total surplus.[30]

In the present paper we focus on replacement strategies. However, compensation incentives can sustain the surplus W for impatience levels below δ^+. This is why compensation incentives are equivalent to transfers of value from the principal to the agent, leaving the total surplus unchanged. Later, we will see that there are perverse effects associated with replacement strategies other than the replacement costs.

Payoff Comparisons - The agent observation of the realized output cannot harm him in payoff terms. Moreover, independently of the information that the principal might hold and given the strategic limitations that we imposed on her behavior, we should expect her payoff to be penalized by the agent's strategic behavior. The following result establishes the relation between the reference payoffs of Definition 24 and the payoffs $v_i^{H,L}$ of Lemma 21 for the interesting case $\delta \in \left[\delta^-, \delta^+\right)$.

Corollary 25 *For* $\delta \in \left[\delta^-, \delta^+\right) \subseteq [0, 1)$ *we have;*

(i) $v_a^{H,L} > v_a^{H,H}$,

[29] A full efficient solution would require $v_a^{H,H} = \pi_a^H$ and $v_p^{H,H} = \pi_p^H$. For that to be possible, the signals had to be perfectly informative in case of high effort, i.e. $p^T = p^S = 0$. Another way to obtain such a result is to employ punishments based on transfers of value between the principal and the agent, see Footnote 23. We choose $p^T > 0$ and $p^S > 0$, to deal with the possibility that nature and/or noisy information may disturb the decision process.

[30] Expression (56) can be obtained along the same lines as in the Proof of Lemma 21. Alternatively, in expression (50), simply replace respectively q^T and π_i^L for p^T and π_i^H.

(ii) $v_p^{H,L} < v_p^{H,H}$.

(iii) Under condition (54), $W > v_a^{H,L} + v_p^{H,L}$ *for all* $\delta \in [0,1)$.

When the agent learns that the monitor type is tolerant and $\delta \in [\delta^-, \delta^+)$, the principal suffers payoff losses. Since the parties in the relation have opposing interests, the result obtained in part (i) justifies the result of part (ii). An extra gain for the agent implies a loss for the principal (not necessarily equivalent).

Part (iii) establishes that high effort in both informational states is surplus superior for any discount level. The extra gains obtained by the agent by deviating to low effort in the informed state, part (i), does not compensate for the payoff losses incurred by the principal, part (ii). This is the case because the gains that the agent obtain are due to a reduction in effort, affecting the generated surplus in an adverse way.

3.4 The Principal Response - Searching for Solutions

In the previous section we have intentionally limited the principal's strategic role to isolate the monitor type revelation effect. We now take a step further towards more realistic scenarios. In addition to the monitoring activities, in an attempt to correct the losses on the principal's payoff, we consider that the principal can strengthen the agent's incentives by costly replace the monitor. Every time the monitor is replaced any learning that the agent has acquired is lost, restoring the uncertainty about the monitor type. The agent is then forced to revise his effort choice.

In the previous section , the information structure was unimportant for the principal, with costly replacement depending on whether she observes the realized output or not leads, to different problems. We consider two potential situations:

(i) Public information with costly replacement, Section 3.5.

(iii) Agent's private information with costly replacement, Section 3.6.

Before exploring these information structures, we consider a solution that has been suggested in Holmström (1982) and Cripps, Mailath and Samuelson (2004), which calls for replacing the monitor in every period. This goal is to establish an upper bound on the principal's payoffs with replacement strategies in the discounting region $\left[\delta^*, \delta^+\right)$.

3.4.1 The Trivial Solution - Costless Replacement

We now discuss the case where there are no replacement costs and the principal is free to substitute the monitors with any desired frequency. The costless replacement case is just a particular case of more general costly structures. Nonetheless, since it corresponds to the opposing limit situation discussed in the previous Section, it deserves exclusive treatment in the sense that it formalizes the principal's best case scenario outside the discount region $\left[\delta^+, 1\right)$. Because of the replacement flexibility, we call the costless scenario, the trivial solution.

Proposition 26 *Independently of the monitoring informational structure and supposing that there are no replacement costs, when $\delta \in \left[\delta^*, \delta^+\right)$, the best strategy for the principal is to replace the monitor in every period and for the agent to supply high effort.*

Conditions (54) and (55) guarantee that $\left[\delta^*, \delta^+\right)$ is nonempty and $\left[\delta^-, \delta^+\right) \subset \left[\delta^*, \delta^+\right)$.

The successive replacement of the monitor disrupts the learning process. The agent cannot profit from the information acquired in one period because the monitor constantly changes. He is never able to update his prior beliefs.

When $\delta \geq \delta^*$, the agent supplies high effort in all periods of the infinitely repeated game. Below δ^*, since the agent incentives are not satisfied, the principal's optimal strategy is to not participate. This scenario is similar to one where the monitor type is known to be θ^E.

It is worth noticing that the repeated replacement of the monitor has the positive effect of sustaining high effort in the discount region $\left[\delta^*, \delta^-\right)$. Until now this was not possible (see Proposition 23).

Denote v_i^{H*} as the payoff that the principal and the agent obtain under permanent costless replacement. Such a payoff can be sustained in the discount region $[\delta^*, 1)$. However, if $\delta \in \left[\delta^+, 1\right)$, the market compensation is enough to sustain incentives for high effort in any informational state. In this case the principal should never replace the monitor.

Even though the permanent replacement solution is able to discipline the agent, it surprisingly achieves payoffs below the reference measure.

Proposition 27 *Independently of the monitoring informational structure and supposing that there are no replacement costs, we have* $v_i^{H,H} > v_i^{H*}$ *for all* $\delta \in (0,1)$ *and* $i \in \{a,p\}$.

This result highlights the negative side of replacement strategies. To understand this result, notice that when incentives for high effort are met, the tolerant monitor is preferred by both the agent and the principal. The reason is that this type incurs less often in mistaken punishments. The permanent replacement of the monitor provides the agent with incentives for high effort for all $\delta \in [\delta^*, 1)$. However, the parties cannot benefit from the extra gains associated with the potential revelation of a tolerant monitor. For that reason, the payoffs under permanent replacement are bounded away from the reference payoffs.

In other words, when $\delta \in \left[\delta^+, 1\right)$, the market compensation is sufficient to provide incentives. The principal should not replace the monitor, she can thus benefit from the potential revelation of a tolerant monitor and obtain $v_p^{H,H}$.[31]

However, below δ^+ the agent incentives for high effort in the informed state collapse. The potential benefit from the revelation of a tolerant monitor disappear. The principal is better off replacing the monitor in every period, keeping the uncertainty about the monitor type always alive.

The results obtained until now lead us to the following conclusion about the principal's chances of recovering the losses incurred due to the revelation of a tolerant monitor when employing replacement strategies.

Corollary 28 *Independently of the monitoring structure, with replacement strategies we have the following bounds on players' payoffs:*

(i) $v_i = v_i^{H,H}$ *for* $i \in \{a,p\}$, *when* $\delta \in \left[\delta^+, 1\right)$.

[31] Care should be taken when comparing v_i^{H*} with $v_i^{H,H}$. The latter is a reference value and can only be sustained in the discount region $\left[\delta^+, 1\right)$, otherwise incentives collapse. The former can be sustained in the interval $[\delta^*, 1)$. Strictly speaking, these payoffs can only be compared for $\delta \in \left[\delta^+, 1\right)$. In the region $\left[\delta^*, \delta^+\right)$ the value v_i^{H*} can be sustained, while $v_i^{H,H}$ is just a reference, see Propositions 23 and 26.

(ii) $v_p \in \left[v_p^{H,L}, v_p^{H*}\right]$ *and* $v_a \in \left[v_a^{H*}, v_a^{H,L}\right]$, *when* $\delta \in \left[\delta^-, \delta^+\right)$.

(iii) $v_p \in \left[\underline{v}_p, v_p^{H*}\right]$ *and* $v_a \in \underline{v}_a \cup \left[v_a^{H*}, v_a^{H,L}\right]$, *when* $\delta \in \left[\delta^*, \delta^-\right)$.

In general terms in any information structure, the principal and the agent cannot obtain a payoff above v_p^{H*} and $v_a^{H,L}$ respectively, the exception is the case $\delta \in \left[\delta^+, 1\right)$. The agent cannot get less than v_a^{H*} unless the costs required to keep the agent with incentives are so high that the principal prefers not to trade. This situation only occurs in the region $\left[\delta^*, \delta^-\right)$ because the principal's choices are constrained by the need to provide the agent with incentives at least in the uninformed state. Otherwise, for $\delta \in \left[\delta^-, \delta^+\right)$, the principal can always guarantee at least $v_p^{H,L}$ by never replacing the monitor.

The result reinforces the message of Proposition 27; we cannot rely on replacement strategies to reach the most efficient outcome. However, these strategies can provide a partial solution to the monitor revelation problem in the discounting region $\left[\delta^*, \delta^+\right)$. For that reason they are worth studying.

Finally, we acknowledge the difficulty of motivating situations with permanent replacement. Nonetheless, the trivial solution seems to fit with the behavior of some managers in specific situations. To be more concrete, consider a monitor that in some days presents a good mood , similar to a type θ^T behavior, while on others days he presents a bad mood, similar to a type θ^S behavior. The subordinates are then unable to identify his state on a particular day. In this case, the manager is using a behavioral strategy, randomizing over the mood θ^S with probability β and the mood θ^T with probability $1-\beta$ in each period. While intuition may support the existence of strategic behavior of this kind in principal-agent relations, further research should verify the validity of such an assertion.

3.5 Public Monitoring with Costly Replacement

In most economic problems, the assumption of free replacement is hard to sustain. We now consider a more interesting scenario where the principal pays a fixed cost $k \geq 0$ every time she decides to replace the existing monitor. These are organizational costs due to adaptation, learning, and/or mandatory firing costs.

The realized output is publicly observed by both the principal and the agent. This information structure captures situations where the principal delegates the monitoring task but, at the same time, keeps track of the realized output. We look at the monitor task as being more complex than simply observing output realizations, he also provides support and assists the agent. Without the monitor no trade would be possible.

Alternatively, we can think that the monitor sends the principal a report at the end of each period with the realized output and the action taken , i.e. the monitor's personal interpretation of the observed signal.

If the true monitor type is tolerant, the principal learns it at the same time as the agent. For that reason it is not rational for the principal to replace the monitor before the occurrence of a revealing signal.

A replacement strategy is then a decision to substitute the existing monitor, $n = 0, 1, 2, ...$ periods after a revealing signal has been observed. Intuitively, when $n = 0$ no learning is possible, the monitor is replaced as soon as an informative signal is observed. This is the highest (rational) replacement frequency and consequently the one with highest total cost. As these costs increase, it might be better for the principal to choose $n = 1$, i.e. to substitute the monitor one period after a revealing signal is observed. The total costs decrease due to a decrease in the replacement frequency, but if $\delta \in \left[\delta^*, \delta^+\right)$, the agent can benefit during one period from learning that the monitor is tolerant. The extreme case is when the monitor is never substituted, i.e. $n \to \infty$. In this case, the principal pays no replacement costs but the agent benefits from the potential revelation of the monitor for the rest of the relation.

The principal faces a trade off between frequent replacement, i.e. right after the monitor revelation, with larger costs but no learning, or less intensive replacement with lower costs but allowing potential deviations from high effort.

Denote $v_{i,k,n}^{0,1}$ as the expected normalized value of the relation for $i \in \{a, p\}$ when the principal replaces the monitor $n \in \mathbb{N}_0 \equiv \{0, 1, 2, ...\}$ periods after a revealing signal, paying the cost $k \geq 0$ per replacement, and the agent effort choice in each informational state is

$\{e^0, e^1\}$.

Lemma 29 *Suppose that the monitoring is public, $k \geq 0$, the principal replaces the monitor $n \in \mathbb{N}_0$ periods after a revealing signal, and the agent chooses $\{e^0, e^1\} = \{e^H, e^L\}$. The infinitely repeated normalized expected payoff for player $i \in \{a, p\}$ is*

$$v_{i,k,n}^{H,L} = \frac{(1-\delta)\pi_i^H + \delta p^E \underline{v}_i + \delta(1-\beta) r \left((1-\delta)\pi_i^L + \delta q^T \underline{v}_i\right) \frac{1-\delta^n(1-q^T)^n}{1-\delta(1-q^T)} - k(1-\delta) D}{1 - \delta\left(1 - p^E - (1-\beta) r \left(1 - \delta^n (1-q^T)^n\right)\right)} + k(1-\delta), \tag{57}$$

where $D \equiv 1 - \delta\left(1 - p^E - (1-\beta) r\right)$, and with $k = 0$ when $i = a$.

The replacement cost k is exclusively incurred by the principal, for that reason when $i = a$ we must set $k = 0$.

The expression (57) incorporates the agent's optimal strategic behavior (best response) for a given principal replacement choice n. Such a construction reduces the computation of the sequential equilibrium of the infinitely repeated game to an optimization problem from the principal's point of view.

The following properties of expression (57) are worth to notice. When $n \to \infty$ we obtain the expression $v_i^{H,L}$ of Lemma 21, i.e.

$$v_{i,k,n}^{H,L} \to v_i^{H,L} = \frac{(1-\delta)\pi_i^H + \delta p^E \underline{v}_i}{D} + \delta(1-\beta)\frac{r}{D}\frac{(1-\delta)\pi_i^L + \delta q^T \underline{v}_i}{1-\delta(1-q^T)}. \tag{58}$$

This is the no replacement case. Similarly when $n = 0$, we obtain

$$v_{i,k,0}^{H,L} = \frac{(1-\delta)\pi_i^H + \delta p^E \underline{v}_i - k\delta(1-\delta)(1-\beta) r}{1-\delta(1-p^E)}, \tag{59}$$

which for $k = 0$ equals expression v_i^{H*}. Notice that with public monitoring and costless replacement, the case where the monitor is replaced after a revealing signal and the case where the monitor is replaced in every period are equivalent. In either situation the agent cannot benefit from learning.

Denote $v_{i,k,n}^{H,H}$ as the value of the relation when, in expression (57), we replace π_i^L and q^T by π_i^H and p^T respectively, and denote $v_{i,k,n}^{L,L}$ as the case where π_i^H, r and p^E are replaced by π_i^L, s and q^E respectively. Similar to Section 3.3, since replacement costs do not enter into the agent's payoff function, the agent's behavior is determined by the discount thresholds that solve $v^{L,L}_{a,0,n} = v^{H,L}_{a,0,n}$ and $v^{H,L}_{a,0,n} = v^{H,H}_{a,0,n}$ which we denote by δ^-_n and δ^+_n respectively.

When the principal replaces the monitor $n \in \mathbb{N}_0$ periods after a revealing signal, if $\delta \geq \delta^+_n$ the agent provides high effort independently of the informational state. On the other hand, for $\delta < \delta^-_n$, low effort is chosen independently of the informational state. Between these two discounting regions, i.e. $\delta \in \left[\delta^-_n, \delta^+_n\right)$, high effort is chosen in the uninformed state and low effort is chosen for n periods, while the tolerant monitor is not replaced after it has been revealed.

In order to conciliate the principal replacement choices with the agent incentives, it is important to understand how δ^-_n and δ^+_n change with n. We start with the latter.

In the neighborhood of δ^+_n, the agent has a dominant strategy to supply high effort while uninformed. The question is, how does his behavior in the informed state change with n. The larger n is, the more it increases the agent's gains in the informed state; since the monitor is tolerant, the relation is expected to last for a larger number of periods. Low effort in this state increases the immediate expected gains but reduces the life expectancy of the relation due to a higher exposure to punishment. As n gets larger the latter effect becomes more important than the former, favoring high effort behavior. Consequently, we must have $\delta^+_{n+1} \leq \delta^+_n$ for all $n \in \mathbb{N}$, or more generally

$$\delta^+ = \delta^+_\infty \leq ...... \leq \delta^+_{n+1} \leq \delta^+_n \leq ... \leq \delta^+_1. \tag{60}$$

The cut-off value δ^+_n is always above δ^+ for all $n \in \mathbb{N}$.[32] This observation supports the results obtained in the previous Sections; when the agent discounts more than δ^+, the principal is better off never replacing the monitor.

To understand the behavior of δ^-_n with respect to n, we start by noticing that in the neighborhood of δ^-_n, low effort in the informed state is a dominant strategy for the agent. What is not clear is how the agent's behavior in the uninformed state changes with n. A deviation to low effort in the uninformed state increases the likelihood of punishment,

[32] Notice also that when $n = 0$, there is no informed state; for that reason δ^+_0 is not defined.

but it also accelerates the potential revelation of a tolerant monitor, because $s > r$. The latter effect is stronger the larger n is due to the larger "revelation prize" in the informed state. The former effect is not affected by variations in n. Consequently, a decrease in n reduces the importance of the latter effect by reducing the "revelation prize", which makes for impatient agents with less incentives to deviate in the uninformed state. Then, we must have $\delta_n^- \leq \delta_{n+1}^-$ for all $n \in \mathbb{N}_0$, i.e.

$$\delta^* = \delta_0^- \leq ... \leq \delta_n^- \leq \delta_{n+1}^- \leq ... \leq \delta_\infty^- = \delta^-. \tag{61}$$

The two extreme values of this sequence can be derived using the limit cases (58) and (59) for $i = a$ and $k = 0$, or as in Sections 3.3 and 3.4 respectively. For $\delta < \delta^*$, the payoffs of the first periods of the relation become more important and, for that reason, low effort becomes a dominant strategy, disregarding the increased punishment likelihood.

As a summary of the preceding discussion, we have the following relation between sets $\left[\delta^-, \delta^+\right) \subseteq \left[\delta_n^-, \delta_n^+\right) \subseteq [\delta^*, 1)$ for all $n \in \mathbb{N}_0$.[33]

Lemma 29 characterizes the agent's strategic behavior as a function of the principal replacement choice n. We now need to find the principal's optimal replacement strategy that maximizes her payoff constrained by the associated costs and the agent's incentives, which vary with δ and n.

Denote k^{PCn} as the cut-off cost value below which the principal's participation is guaranteed when she replaces the monitor n periods after the observation of a revealing public signal and $\delta \in \left[\delta^*, \delta^-\right)$. Let k^{cut} be a reference threshold cost that solves $v_{p,k,n}^{H,L} = v_{p,k,n+1}^{H,L}$. Unconstrained by any incentives, when $k < k^{cut}$, $n = 0$ is optimal because $v_{p,k,n}^{H,L} > v_{p,k,n+1}^{H,L}$ for all n. Otherwise $n \to \infty$ is the optimal choice. In the discount region $\left[\delta^-, \delta^+\right)$ the value $k^{cut} = k^{RC}$ is the replacement threshold above which the principal prefers to never replace the monitor. Recall that when the agent discounts on this region, the principal always has the option of never replacing the monitor, thus securing a payoff of $v_p^{H,L} > \underline{v}_p$.

[33] When n gets large we obtain higher order polynomials. The problem becomes untractable and the discount thresholds δ_n^- and δ_n^+ have to be computed numerically. However, since we know how δ_n^- and δ_n^+ vary with n, this is enough for our proposes.

Proposition 30 *Suppose that the information is public and $k \geq 0$, the principal's best strategy:*

(i) When $\delta \in \left[\delta^{-}, \delta^{+}\right)$, is to choose $n = 0$ for $k \in \left[0, k^{cut}\right)$ and $n \to \infty$ for $k \in \left[k^{cut}, \infty\right)$. Where

$$k^{cut} \equiv \frac{\pi_p^H - \underline{v}_p}{1 - \delta\left(1 - p^T - r\right)}. \tag{62}$$

(ii) When $\delta \in \left[\delta_n^{-}, \delta_{n+1}^{-}\right) \subset \left[\delta^{*}, \delta^{-}\right)$, is to choose $n = 0$ for $k \in \left[0, k^{cut}\right)$, n for $k \in \left[k^{cut}, k^{PCn}\right)$, and no trade otherwise. Where

$$k^{PCn} \equiv \frac{\pi_p^H - \underline{v}_p}{\delta\left(1 - \beta\right) r \delta^n \left(1 - q^T\right)^n}. \tag{63}$$

When the replacement costs are sufficiently low, the principal's best strategy is to replace the monitor after a revealing signal, not allowing the agent to learn. This is true providing $k < k^{cut}$ and $\delta \in \left[\delta^{*}, \delta^{+}\right)$. The choice $n = 0$ guarantees high effort in both informational states. On the other hand, if replacement costs are large, i.e. $k \geq k^{cut}$ and $\delta \in \left[\delta^{-}, \delta^{+}\right)$, it is better to never replace the monitor, i.e. $n \to \infty$. However, when $\delta \in \left[\delta_n^{-}, \delta_{n+1}^{-}\right) \subset \left[\delta^{*}, \delta^{-}\right)$ and $k^{cut} \leq k < k^{PCn}$, in order to keep the agent with incentives for high effort in the uninformed state, the principal must replace the monitor n periods after a revealing signal. In other words, to keep the agent with incentives in the uninformed state the principal must allow the agent to benefit from learning during n periods. However, such a demand might be too costly for the principal, i.e. $k \geq k^{PCn}$, in which case she prefers to not trade and get the outside option.

Recall that the optimal behavior for $\delta \in \left[\delta^{+}, 1\right)$ by Proposition 23 is to choose $n \to \infty$.

From (62) we can see that a choice $n = 0$ is favoured; when the difference between the principal's stage game gains and the outside option is larger, when the agent is more patient or the punishment and learning probabilities are lower. The effect of these variables is intuitive and not surprising. The value of k^{PCn} in (63) is affected in the same way by these variables, but also increases; the larger n becomes, the smaller the proportion of tolerant monitors becomes or larger the punishment probability of a tolerant monitor in the case of low effort.

In the limit, as $n \to \infty$, we obtain expression (50) as a particular case of (57). This choice cannot improve the principal's payoff. The strategic limited setting of Section 3.3 provides a lower bound on the payoffs that the principal can obtain, i.e. $v_p^{H,L}$ for $\delta \in [\delta^-, \delta^+)$ and $\underline{v}_p$ for $\delta \in [\delta^*, \delta^-)$. When the optimal choice takes a finite number of periods and the principal wants to trade, it must be because the replacement costs are sufficiently small and allows for payoff improvements. The following result formalizes this intuition.

Corollary 31 *Suppose that the information is public and $k \geq 0$. If $\delta \in [\delta^-, \delta^+)$ and the optimal choice is $n = 0$ the principal improves her payoffs w.r.t. $v_p^{H,L}$. If $\delta \in [\delta^*, \delta^-)$, the optimal choice n is finite and there is trade, then the principal improves her payoffs w.r.t. $\underline{v}_p$.*

The principal's payoff losses associated with the revelation of the monitor type can be partially recovered when the replacement costs are not too high. These strategies are particularly powerful under public monitoring because the principal enjoys a great replacement precision. However, these strategies require some extra destruction of value due to the replacement costs;[34] consequently, the principal's expected payoff is bounded from v_p^{H*} in the interval $[\delta^*, \delta^+)$, see Corollary 28. The higher the replacement costs, further down the principal's payoff is pushed.

Figure 10 illustrates the value of $v_{p,k,n}^{H,L}$ for the cases where $k = 1$ and $k = 30$ when $\delta \in [\delta^-, \delta^+)$. Since $k^{cut} = 20.24$, in the former case, $n = 0$ is optimal, while in the latter $n \to \infty$ is the optimal replacement choice. Notice also, how both functions converge to $v_p^{H,L}$ as $n \to \infty$. The value of v_p^{H*} and the reference value $v_p^{H,H}$ are also shown.

3.5.1 Public Information with Compensation Incentives

We complete this section by discussing the potential of compensation schemes as an instrument to strengthen the agent's incentives for high effort.

It is common in the incentives literature, under the usual constraints, to allow the principal to freely set the compensation. Translated to our setting, the principal would choose a compensation in the informed state and a compensation in the uninformed state. This

[34] Recall that there is also destruction of value due to mistaken punishments on the equilibrium path. Value burned due to replacement cost represents an extra layer in terms of loss in value.

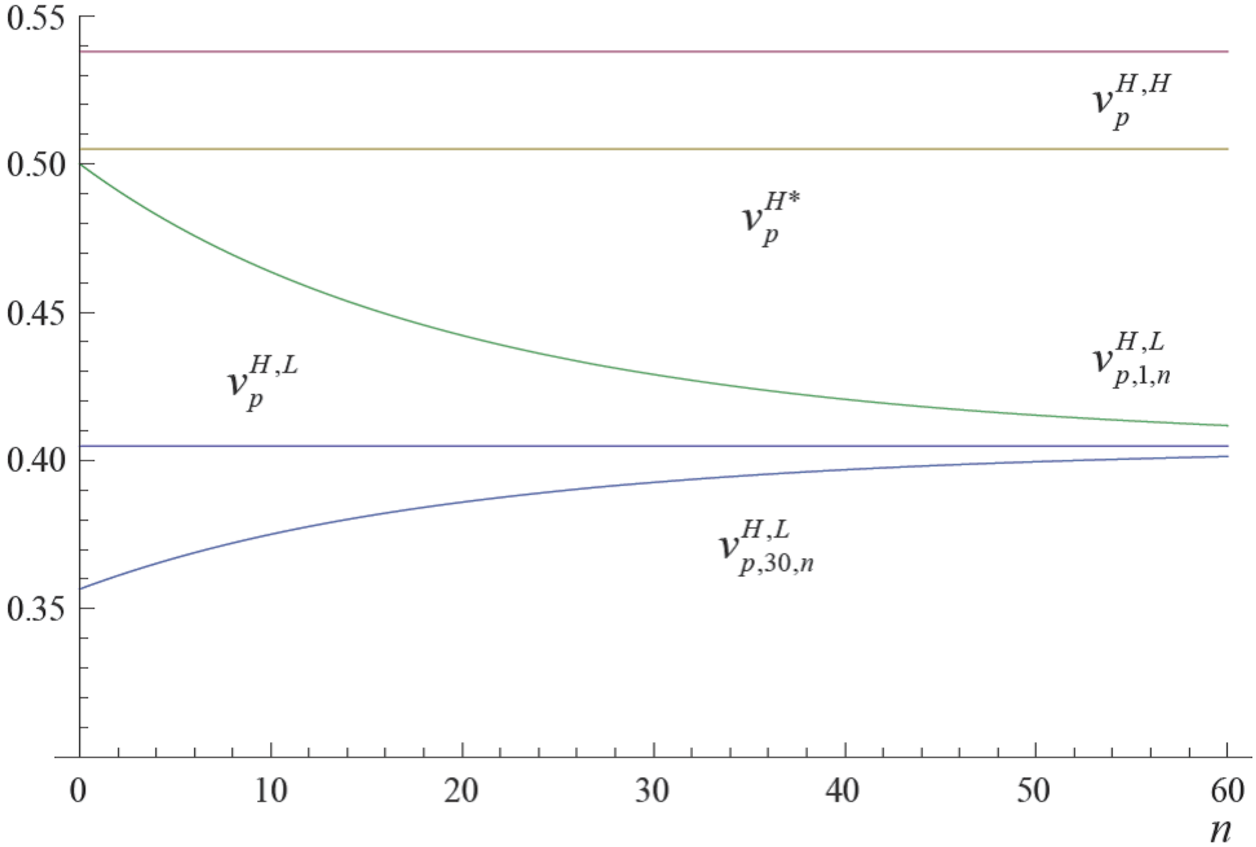

Figure 10: Principal's payoffs under public monitoring for different replacement costs.

case is particularly penalizing for the agent. Since the incentives for high effort in both informational states are related, the principal would find it optimal to offer a low compensation $w^0 < w$ in the uninformed state, retaining a larger fraction of the surplus in the initial periods ("exploitation" state) of the relation, in exchange for a higher compensation $w^1 > w$ after a revealing signal ("reward" state).[35]

Since players discount the future, the initial larger gains become more relevant; for that reason we might observe not only a payoff improvement above the value v_p^{H*}, but also above the reference value $v_p^{H,H}$. Since the total surplus remains constant, such improvements are made at the agent's expense. (see Footnote 37 below)

For this mechanism to work as described, the wage has to be sufficiently flexible. Situations like this one might require some intervention in favor of the agent. Wage restrictions, as we shall discuss, may provide a partial remedy to the problem.

Consider a more realistic scenario where compensations below the market value w cannot be offered.[36] The lower bound restriction w limits the principal's exploitation of the agent in the uninformed state.

To be more concrete, suppose that both the agent and the principal discount the future according to some $\delta \in [\delta^-, \delta^+)$. In this case the principal can offer a higher compensation in the informed state $w^1 > w$, moving δ^+ down towards δ, providing the agent with incentives for high effort in this state.[37] More impatient agents would require a higher w^1 but there is a bound on the compensation that the principal can offer, i.e. $be^H - w^1 \geq \underline{v}_p$, otherwise the principal would prefer not to trade. However, she will still be able to sustain high effort

[35] It represents an alternative explanation that justifies why young employees tend to work harder in the first period of their careers, relaxing later. See Medoff and Abraham (1981) for some early empirical evidence supporting this fact. Paradoxically it is also in the early stages of an individual career that compensations tend to be lower. See also Gibbons and Murphy (1992).

[36] See Taylor (1980) for early empirical observations about wage stickiness. Wage stickiness is particularly strong, even in periods of recession, see Bewley (1999).

[37] There is a second effect. Higher compensation in the informed state helps on the provision of incentives in the uninformed state, since the potential losses associated with an increase in the likelihood of punishment due to a deviation in this stage becomes more important. Consequently, the principal has higher freedom to offer a lower compensation in the uninformed state $w^0 < w$, increasing δ^- up to δ. However, this effect will not be available since the principal is restricted to $w^0 \geq w$. Notice also that if players are very patient, i.e. $\delta \geq \delta^+$, we observe $w^1 \leq w$. Independently of the informational state, the more patient players are, the more surplus the principal appropriates.

in the uninformed state by offering $w^0 \geq w$, pushing δ^- down towards δ.

The increasing compensation path $w^0 \leq w^1$ has empirical support. During a worker's career the salaries tend to increase above the market reference value. In our setting such an event is caused by the perverse effect associated with the agent's learning.

When high effort is always played in equilibrium, we obtain the same total surplus as in the reference scenario. Compensation schemes as described are transfers of value from the principal to the agent. For that reason, it is not clear whether compensation schemes are payoff superior from the principal's perspective to replacement strategies, in particular when the agent is sufficiently impatient.[38] Improvements over the reference value are impossible for the principal, it is the agent that benefits from the higher compensation. Nonetheless, improvements over the payoff v_p^{H*} are possible when the agent is sufficiently patient.

In a multiagent situation, replacement strategies have an amplifying effect. Typically, the same monitor interacts with multiple agents and a specific noisy signal of each agent effort is observed.[39] In this case, once one of the agents has learned that the monitor is tolerant, the principal either decides to increase his compensation or allows him to benefit from his learning. The latter solution leads to relaxation in his performance. The former case leads to a situation where one individual with the same average performance and the same qualifications is paid more than his colleagues. This situation might bring discontentment and a sense of unfairness among the other workers. Moreover, there is also the risk that the monitor type becomes common knowledge due to communication between workers. In order to reestablish the incentives, the principal has to increase the compensation to all the other workers, this solution might be extremely expensive.

Another possibility would be to suggest the replacement of that specific agent or even the whole group. The latter solution seems extremely expensive. In the former case, firing

[38] A clear answer to this questions depends not only on δ but also on the cost k associated with the replacement of the monitor.

[39] This setting is distinct from the one studied in the Seminal work of Holmström (1982b) with moral-hazard-in-teams, where only the team aggregate effort measure is observed. Here, a signal of each agent effort is observed.

without justification is usually more costly than firing when there is evidence of low performance which can be shown in a court of law. Moreover, a sense of injustice may also emerge among the group members.

The preceding example involves behavioral considerations that go beyond the scope of the present paper. However, it might rationalize, for example, why in some sporting activities, after a sequence of bad results, it is the coach that is replaced not the whole team.

3.6 Agent's Private Monitoring with Costly Replacement

We now consider the possibility that the agent and the monitor observe the realized output while the principal does not.

This situation is typical in large corporations where the management and the monitoring functions are separated. The principal only observes an aggregate measure of the full output produced by a particular department or by the whole firm. This measure includes the contributions of all the individuals involved in the production process. The principal cannot disentangle the output of a particular agent from that of the other individuals.[40] It is then the responsibility of lower rank managers (the monitor in our case) to take decisions about a particular individual. The monitor has an informational advantage, for that reason the principal fully delegates the monitoring task.[41] The monitor then reports to the principal when a relevant event has occurred, i.e. an interpretation of low effort.

Since the principal is informed about a punishment event we are able to keep the recursive structure of the problem. To be precise, the setting of this Section is one of private monitoring with communication.[42]

[40] If the principal were able to observe the payoffs associated with the effort choices of a particular agent, she would be able to infer the agent noise signal by looking at her own payoff.

[41] That does not mean that the monitor's work is not object of monitoring. We can assume that a higher hierarchy monitor verifies if the monitor is performing his work according to the standards defined by the corporation. This issue is related with the firm's organization design and boundaries. We refer the reader to Rahman (2009), which suggests some interesting answers to the question; who monitors the monitor?

[42] Kandori (2002) presents a description of the challenges associated with private monitoring. Early folk-theorems for private monitoring with communication were obtained by Compte (1998) and Kandori and Matsushima (1998). In our setting we allow for mistaken punishments and we do not consider transfers of value among the players, for that reason we are always bounded far from full efficiency. More recently, Obara (2009) and Zheng (2008) relaxed some of the assumptions of the early contributions.

Unlike in the public monitoring case, since the principal does not observe the noisy measure of the agent effort, she misses the monitor's revelation process. To be more concrete, suppose the true monitor type is tolerant and the agent receives a revealing signal and consequently supplies low effort. The principal does not know in which moment in time (or even if) this revelation has occurred. Unlike in Section 3.5, this "reference moment" is not available.

Nonetheless, the principal knows the model and all the associated parameters; moreover, she knows what payoffs are due in case of low and high effort. Given her knowledge about the whole problem, the principal has to design a replacement scheme that is optimal given her "blind" position.

For that reason, her replacement choices are always limited, either because they are premature, in the sense that the agent was still uninformed incurring in an unnecessary cost, or because they are late, in the sense that the monitor type was already revealed and the agent is providing low effort. These imprecisions weaken the effectiveness of replacement strategies in private monitoring contexts.[43] However, that does not imply that improvements over the strategic restricted payoff $v_p^{H,L}$ are not possible.

As mentioned in the introductory Section, auditing companies and financial supervision authorities experience a problem with a similar information structure. Auditing companies (the principal) rotate the external auditors (the monitor) on a regular basis.[44] This practice attempts to eliminate what is known in the accounting/auditing jargon as the "familiarity threat" between the client (the agent) and the auditor.[45] Through repeated interaction, the auditor reveals professional and personal characteristics to the client. Learning issues

[43] The same weakness would be present in any other incentives scheme, the difficulty is in the information structure.

[44] The Sarbanes-Oxley Act (enacted on July 30, 2002) in Section 203, requires the lead audit partner and audit review partner (or concurring reviewer) to be rotated every five years on public company audits as well as on audits of issuers.

[45] Most of the literature on incentives studies this practice as a remedy to the breakout of collusive arrangements. Tirole (1986) points out that when the monitoring task is delegated to a third party, problems related with the monitor's conflicting interests might arise. See Laffont and Tirole (1991) and Kofman and Lawarree (1993) for further developments and extensions. In this paper we are not so concerned with delegation effects of this kind.

of this kind favor the occurrence of strategic behaviour from the client.[46]

The auditing company is hardly aware of these facts but knows that they are likely to occur. Given the information structure and the agent's strategic behavior; with which frequency should the auditing company rotate its external auditors? This Section provides an answer to this question.

Now, n is the number of periods a given monitor stays in charge, after being hired and after the first signal realization. ($n+1$ is the actual number of periods that the monitor is employed) For example $n=0$ means that the monitor is replaced every period, i.e. is employed for a single period, while $n=1$ means that the monitor is replaced every second period and so on. Notice the difference in the interpretation of n with respect to the public monitor case of Section 3.5.

Denote $\widetilde{v}_{i,k,n}^{0,1}$ as the expected normalized value of the relation for $i \in \{a,p\}$ when the principal replaces the monitor $n \in \mathbb{N}_0$ periods after have hiring him, paying the replacement cost $k \geq 0$ every time, and the agent effort choice in each informational state is $\left\{e^0, e^1\right\}$.

Lemma 32 *Suppose that the information is the agent's private monitoring, $k \geq 0$, the principal replaces the monitor $n \in \mathbb{N}_0$ periods after hiring him and the agent chooses $\left\{e^0, e^1\right\} = \left\{e^H, e^L\right\}$. The infinitely repeated normalized expected payoff for player $i \in \{a,p\}$ is*

$$\widetilde{v}_{i,k,n}^{H,L} = \frac{\Pi_i^H \frac{1-\delta^{n+1}z^{n+1}}{1-\delta z} + \delta y \Pi_i^L \frac{(z-x)(1-\delta^n z^n)-x(z^n-x^n)\delta^n(1-\delta z)}{(z-x)(1-\delta x)(1-\delta z)} - k(1-\delta)}{1-\delta^{n+1}\frac{(z-x)z^{n+1}+y(z^{n+1}-x^{n+1})}{z-x}} + k(1-\delta), \quad (64)$$

where $x \equiv 1-q^T$, $y \equiv (1-\beta)r$, $z \equiv 1-p^T-r$, $\Pi_i^H \equiv (1-\delta)\pi_i^H + \delta p^E \underline{v}_i$ and $\Pi_i^L \equiv (1-\delta)\pi_i^L + \delta q^T \underline{v}_i$. When $i=a$ we have $k=0$.

Expression (64) has the following asymptotic properties. When we let $n \to \infty$ we obtain $v_i^{H,L}$ as expression (58) in the previous Section. While, if $n=0$, we obtain

$$\widetilde{v}_{i,k,0}^{H,L} = \frac{(1-\delta)\pi_i^H + \delta p^E \underline{v}_i - k\delta(1-\delta)\left(1-p^E\right)}{1-\delta\left(1-p^E\right)}. \quad (65)$$

[46] The "familiarity threat" may also be caused by collusion between the agent and the monitor, but even in this case some prior learning has to occur. A rational dishonest client would not bribe an external auditor without a prior observation of his personal character traits, otherwise he could place himself in a worse situation.

When $k = 0$ we get the expression v_i^{H*}.

Since $1 - p^E > (1 - \beta) r$, the principal's payoff in (65) is smaller than in (59). The difference in payoffs reflects the loss in precision of replacement strategies under private monitoring. In other words, in Section 3.5, the replacement of the monitor was only an issue after the arrival of a revealing signal which was publicly observed, while under the agent's private monitoring, replacement is a possibility from the first period of the relation.

As in Section 3.5, we focus our attention on the discounting interval $\left[\delta^*, \delta^+\right)$. In particular, in discounting region $\left[\delta^*, \delta^-\right)$, we have a sequence of δ_n^- ordered as in (61) of Section 3.5 (for the intuition, the reader is referred to the discussion in that Section). However, under private monitoring δ_n^- is smaller. This is because the interpretation of n under private and public monitoring are different. With public monitoring, n is the number of periods that the tolerant monitor remains in charge after a revealing signal, while under private monitoring, n is the number of periods that a given monitor stays in charge after the first signal. Consequently, for the same n, replacement is more frequent under private than under public monitoring, justifying the difference of δ_n^- under the different information structures.[47]

To better distinguish between both information structures, suppose that a revealing signal occurs at time $t \geq 1$. When the signals are public, the monitor stays in charge for a total of $t+n$ repetitions of the stage game. While, if monitoring is private, the monitor stays in charge for a total of $1 + n$ repetitions of the stage game, and we might have $t \gtrless n + 1$.

Rotation of the monitor under public information is more accurate but the replacement cycle $t+n$ is stochastic. This is the case because the reference (revelation) period t is random, unknown ex-ante but observed ex-post. Replacement strategies under private monitoring are less accurate, since t is random and not known by the principal even ex-post. For that reason, the replacement cycle n is defined and known ex-ante.

Denote $k^{n,n+1}$ as the cut-off point which, infinitesimally below n, is an optimal choice for the principal and, infinitesimally above $n + 1$, is optimal. It is a transition threshold

[47] On the other hand, δ_n^+ under private monitoring is larger than δ_n^+ under public monitoring. Which is consistent with the observation made about δ_n^-. An equal ordered sequence of δ_n^+ as in (60) is obtained.

between replacement choices. For fixed k, when $k^{n-1,n} \leq k \leq k^{n,n+1}$, n is an unconstrained optimal choice. Let $k^{\infty} \equiv \lim_{n \to \infty} k^{n,n+1}$ be the value above which the principal's optimal unconstrained choice is to never replace the monitor. We say unconstrained because we are not considering any incentives constraint.

For $\delta \in [\delta^{-}, \delta^{+})$, let k^{RCn} denote the per replacement cost threshold below which the principal's optimal choice can be n. When $k^{RC\infty} \leq k$, the choice $n \to \infty$ must be optimal. Similarly, for $\delta \in [\delta^{*}, \delta^{-})$, let k^{PCn} denote the participation condition below which n can be optimal. To keep the text clean, the functional form of each of these objects can be found in the Proof of the following result.

Proposition 33 *Suppose that the information is the agent's private monitoring and $k \geq 0$. Let τ denote the principal optimal choice.*

(i) For $\delta \in [\delta^{-}, \delta^{+})$: If $k < k^{\infty}$, then $\tau = \inf \{n, n+1, ... : k < k^{RC\tau}\}$. Otherwise, i.e. $k^{\infty} = k^{RC\infty} \leq k$, $\tau \to \infty$.

(ii) For $\delta \in [\delta_{m}^{-}, \delta_{m+1}^{-}) \subseteq [\delta^{*}, \delta^{-})$ and $n \leq m$: If $k < k^{\infty}$, then $\tau = \inf \{n, n+1, ..., m : k < k^{PC\tau}\}$, while if $k^{\infty} \leq k < k^{PCm}$, then $\tau = m$. Otherwise, i.e. $k \geq k^{PCm}$, there is no trade.

(iii) For $\delta \in [\delta_{m}^{-}, \delta_{m+1}^{-}) \subseteq [\delta^{*}, \delta^{-})$ and $n > m$: If $k < k^{PCm}$, then $\tau = m$. Otherwise, i.e. $k \geq k^{PCm}$, there is no trade.

The principal participation and consequent employment of replacement strategies is guaranteed when $\widetilde{v}_{p,k,n}^{H,L} > \underline{v}_{p}$ for $\delta \in [\delta^{*}, \delta^{-})$ and $\widetilde{v}_{p,k,n}^{H,L} > v_{p}^{H,L}$ for $\delta \in [\delta^{-}, \delta^{+})$. The agent incentives for high effort in the uninformed state are guaranteed for $\delta \geq \delta_{n}^{-}$. When we join together these restrictions with the unconstrained optimal choice we obtain the equilibrium strategies for the agent and the principal.

The agent is rational and knows the replacement cycle. Consequently, when the principal replaces the monitor, the agent also shifts from low to high effort. Then he waits for the occurrence of a potential revealing signal before the next monitor replacement, in order to enjoy the remaining period providing low effort. The principal's strategy takes into account this strategic behavior.

In the discounting region $[\delta^-, \delta^+)$ the principal's best replacement strategy is no longer an exclusive choice between $n = 0$ and $n \to \infty$, as we found in Proposition 30. Now the equilibrium is more sensitive to the value k. In fact any n can be optimal, for that reason high effort in any informational state is only possible for $k < k^{0,1}$.[48] The reason is the trade-off between more frequent replacement, more costly but less likely to allow the agent to learn, and less frequent replacement, cheaper but with a higher probability that the principal's payoffs will be penalized by the revelation of a tolerant monitor.

When $\delta \in [\delta^-_m, \delta^-_{m+1}) \subseteq [\delta^*, \delta^-)$ and the unconstrained optimal choice $n \leq m$, the principal can choose n if $k < k^{PCn}$. Otherwise, since k^{PCn} is strictly increasing with n, she has to move way from the optimum in the direction of m. However, it might still happen that $k \geq k^{PCm}$, in which case the principal should not trade because replacement costs are too high. A choice above m is cheaper but does not provide the agent with incentives.

When $\delta \in [\delta^-_m, \delta^-_{m+1}) \subseteq [\delta^*, \delta^-)$ and the unconstrained optimal choice $n > m$, the principal must increase the replacement frequency, moving way from the optimum, down towards m to provide the agent with incentives in the uninformed state. However, if $k \geq k^{PCm}$ the provision of incentives is too costly and the principal should not enter into the relation.

When $n \to \infty$ we obtain expression (58) as a particular case of (64). It is better to allow the agent to provide low effort in the informed state than to make any costly replacement, see Corollary 28. Nonetheless, for sufficiently low replacement costs, it is possible to make a payoff improvement using replacement strategies even without observing the realized output. The following result is in everything similar to Corollary 31.

Corollary 34 *Any equilibrium of Proposition 33 with a finite choice n and trade, improves the principal's payoff* $\widetilde{v}^{H,L}_{p,k,n}$ *over the payoffs associated with Proposition 23, i.e.* $v^{H,L}_p$ *for* $\delta \in [\delta^-, \delta^+)$ *and* $\underline{v}_p$ *for* $\delta \in [\delta^*, \delta^-)$.

For $\delta \in [\delta^*, \delta^-)$ we require $\widetilde{v}^{H,L}_{p,k,n} > \underline{v}_p$, which is the participation or a replacement constraint that can only be satisfied if n is finite. When $n \to \infty$ no trade is an equilibrium. For

[48] It can be shown that $k^{0,1}$ is a very small positive number.

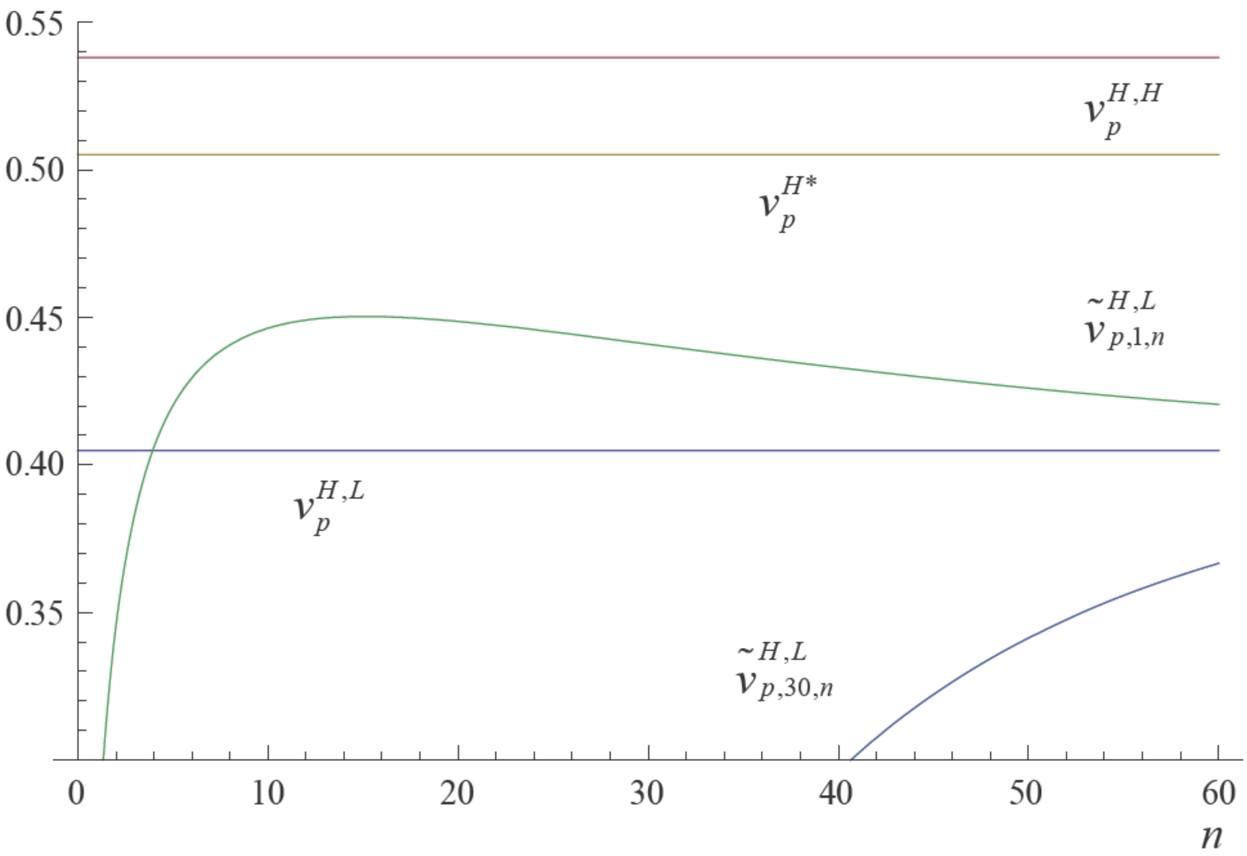

Figure 11: Principal's payoffs under private monitoring for different replacement costs.

$\delta \in [\delta^-, \delta^+)$ we require that $\tilde{v}_{p,k,n}^{H,L} > v_p^{H,L}$, which is only possible if replacement strategies are employed.[49] If $n \to \infty$ we have $\tilde{v}_{p,k,n}^{H,L} \to v_p^{H,L}$, i.e. the worst case scenario for the principal, that she can always guarantee by not replacing the monitor.

Figure 11 illustrates the value of $\tilde{v}_{p,k,n}^{H,L}$ for the cases where $k = 1$ and $k = 30$ when $\delta \in [\delta^-, \delta^+)$. In the former case $k^{13,14} = 0.93 < k^{14,15} = 1.04$ and $k^{RC14} = 2.89$, then $n = 14$ is the optimal choice. Since $k^{\infty} = 20.24$ when $k = 30$, $n \to \infty$ is optimal. Both functions converge to $v_p^{H,L}$ as $n \to \infty$.

3.6.1 Agent's Private Information with Compensation Incentives

We now comment on the possibility of compensation based incentives. As in Subsection 3.5.1 we assume that the agent will not accept to work for less than the market wage. The

[49] For that reason we call it a replacement constraint. Since $v_p^{H,L}$ is always larger than $\underline{v}_p$ we cannot talk about a participation constraint in the strict sense.

principal is allowed to raise the compensation if she considers it convenient but cannot decrease it.

To understand how compensation incentives can be used in a private monitor setting, consider the following strategy. During the first $n \in \mathbb{N}_0$ periods of the relation, the principal pays the compensation $w^0 \geq w$. In the following periods, she switches to the compensation $w^1 \geq w$, with $w^0 < w^1$. The idea is to offer a lower compensation sufficient to provides the agent with incentives for high effort in the early periods, where it is less likely that the monitor type has been revealed to be tolerant, and then adjust to a higher compensation when it is more likely that the monitor has been revealed.

Now, together with the uninformed and the informed states of the relation, we have the low compensation stage and the high compensation stage. A strategy for the principal is a choice of w^0, w^1 and the stage separating period n. Again we might have $n = 0$, i.e. the relation starts in the high compensation stage, or $n \to \infty$, i.e. the relation remains in the low compensation stage forever, or an optimal intermediate choice of n, i.e. the relation passes through both stages.

The informational disadvantage of the principal with respect to the agent will necessarily reflects in a lower payoff for the former when compared with the case where the realized output is publicly monitored. Again, we expect mixed superiority of replacement strategies with respect to compensation incentives. The latter must be stronger when the agent is more patient and/or replacement costs are sufficiently large.

The two stage compensation scheme discussed here is similar to the existing one in the public sector. There is a distance, not only physical but also in monitoring terms, between the central authority and the lower hierarchical levels. The performance evaluation and the functioning of the associated public office is usually based on general reports. For that reason promotions, measured in compensation benefits, are usually independent of the performance but rather depend on years accumulated in service.

3.7 *Final Comments*

In many economic situations of interest managers have the necessity to delegate some of their tasks - in this model, the monitoring activities, which are crucial for the regular functioning and expansion of their businesses. The degree of delegation in some sense determines the subsequent information structure. Full delegation leads to an information structure similar to the agent private monitoring case discussed in this paper. Partial delegation is closer to public monitoring information structures. The present paper provides some results, about the optimal strategic behavior from the principal's perspective, for dealing with the negative effects associated with the revelation of specific organizational aspect which might be the object of adverse strategic behavior. We choose the monitor type to be the unknown piece of information that a potentially strategic agent might take advantage of once informed, but the spectrum of situations with similar characteristics is larger.

The principal usually has more freedom in the choice of the incentives schemes, in this paper we focus on replacement strategies. As mentioned before, the revelation of the player's type through repeated interaction is not a new finding. However, the way such a problem is modeled in this paper is novel.

Many questions are left open. For example, a clearer connection with the existing theories in multiple and common agency, renegotiation-proof, incomplete contacts, private evaluations, information sharing, etc. Compensation incentives were discussed but not formalized.

We also did not cover all potential information structures, for example the possibility of the realized output being exclusively observed by the principal (and/or the monitor) or when the principal holds prior private information about the monitor type. These cases capture situations where the principal has access to relevant information that for some reason, intentionally or not, it is blocked to the agent. We expect the principal's informational advantage to help her in achieving higher payoffs. In the former possibility, the principal can be more efficient in her replacement choices, in particular if the tolerant type is preferred. However, the principal cannot replace the monitor successively until a tolerant type appears because that behavior would reveal the monitor type to the agent. There is here a trade off between replacement costs and payoff gains with a tolerant monitor. An optimal strategy for the principal must require some degree of randomization between replacement choices.

The contrast between the agent and monitor replacement is also an interesting point. Another possibility is to allow the monitor to play a strategic role or even to remove him and consider a strategic principal with an unknown type to the agent. This lead us to dynamic incentive problems of incomplete information, typically harder to handle but very rich in strategic terms.

Also interesting, but from a different perspective, is the introduction of new ingredients into the problem, and more empirical and experimental work on the subject are the next steps towards a better understanding of this type of revelation problems. Such research should also provide us with recommendations on how we could implement the proposed solutions in our organizations.

3.8 References

1. Abreu, D., D. Pearce and E. Stacchetti (1986). "Optimal Cartel Equilibria with Imperfect Monitoring," Journal of Economic Theory, 39, 251-269.

2. Abreu, D., D. Pearce and E. Stacchetti (1990). "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." Econometrica, 58, 1041-1063.

3. Alchian, A. and H. Demsetz (1972). "Production, Information Costs, and Economic Organization." American Economic Review, 62, 777-795.

4. Baker, G., R. Gibbons and K. J. Murphy (1994). "Subjective Performance Measures in Optimal Incentive Contracts." Quarterly Journal of Economics, 109, 1125-1156.

5. Baker, G., R. Gibbons and K. J. Murphy (2002). "Relational Contracts and the Theory of the Firm." Quarterly Journal of Economics, 117, 39–84.

6. Bewley, T. F. (1999) Why Wages Don't Fall During a Recession. Harvard University Press.

7. Bolton P., M. Dewatripont (2005) Contract Theory. MIT Press, Cambridge, MA.

8. Bull, C. (1987). "The Existence of Self-Enforcing Implicit Contracts." Quarterly Journal of Economics, 102, 147–159.

9. Cole H., J. Dow, W. B. English (1995). "Default, Settlement, and Signalling: Lending Resumption in a Reputational Model of Sovereign Debt." International Economic Review, 36, 365-385.

10. Compte, O. (1998). "Communication in Repeated Games with Imperfect Private Monitoring," Econometrica, 66, 597-626.

11. Cripps, M. W., J. C. Ely, G. J. Mailath, and L. Samuelson (2008). "Common Learning." Econometrica, 76, 909-933.

12. Cripps, M. W., G. J. Mailath, and L. Samuelson (2007). "Disappearing Private Reputations in Long-Run Relationships." Journal of Economic Theory, 134, 287–316.

13. Cripps, M. W., G. J. Mailath, and L. Samuelson (2004). "Imperfect Monitoring and Impermanent Reputations." Econometrica, 72, 407–432.

14. Ferguson, T. (1989). "Who solved the secretary problem?" Statistical science, 4, 282-296.

15. Fuchs, W. (2007). "Contracting with Repeated Moral Hazard and Private Evaluations," American Economic Review, 97, 1432-1448.

16. Fudenberg, D., D. Levine (1994). "Efficiency and Observability with Long-Run and Short-Run Players." Journal of Economic Theory, 62, 103–135.

17. Fudenberg, D., D. Levine and E. Maskin (1994). "The Folk Theorem with Imperfect Public Information." Econometrica, 62, 997-1040.

18. Fudenberg, D. and E. Maskin (1986). "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," Econometrica, 54, 533-554.

19. Fudenberg, D. and J. Tirole (1991) Game Theory, MIT Press, Cambridge, MA.

20. Fudenberg, D. and J. Tirole (1986), "A "Signal-Jamming" Theory of Predation." Rand Journal of Economics, 17, 366–376.

21. Freeman, P. (1983). "The secretary problem and its extensions: A review." International Statistical Review, 51, 189-206.

22. Gibbons, R. and K. J. Murphy (1992). "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence." The Journal of Political Economy, 100, 468-505.

23. Gossner, O. and T. Tomala (2009). "Repeated Games with Complete Information", in Meyers, Robert (Ed.) Encyclopedia of Complexity and Systems Science, pp 7616–7630. Springer New York.

24. Holmström, B. (1982a). "Managerial Incentive Problems - A Dynamic Perspective", in Essays in Economics and Management in Honor of Lars Wahlbeck. Helsinki: Swedish School of Economics. (See also Review of Economic Studies, (January 1999), 66, 169-182)

25. Holmström, B. (1982b). "Moral Hazard in Teams.", Bell Journal of Economics, 10, 74-91.

26. Kandori, M. (2002). "Introduction to Repeated Games with Private Monitoring," Journal of Economic Theory, 102, 1-15.

27. Kandori, M and H. Matsushima (1998) "Private Observation, Communication and Collusion," Econometrica, 66, 627-652.

28. Kennan J. and R. Wilson (1989). "Strategic Bargaining Models and Interpretation of Strike Data," Journal of Applied Econometrics, 4, Supplement: Special Issue on Topics in Applied Econometrics, S87-S130.

29. Klein B., and K. B. Leffler (1981). "The Role of Market Forces in Assuring Contractual Performance." The Journal of Political Economy, 89, 615-641.

30. Kofman, F. and J. Lawarree (1993), "Collusion in Hierarchical Agency", Econometrica, 61, 629-656.

31. Laffont, J. and J. Tirole (1991), "The Politics of Government Decision-Making: A Theory of Regulatory Capture", Quarterly Journal of Economics, 106, 1089-1127.

32. Leibenstein, H. (1987) Inside the Firm: The Inefficiencies of Hierarchy, Harvard University Press, Cambridge, MA.

33. Levin J. (2003). "Relational Incentive Contracts." American Economic Review, 93, 835-857.

34. MacLeod W. B., (2003). "Optimal contracting with subjective evaluation." American Economic Review, 93(1), 216-240.

35. MacLeod W. B., and J. M. Malcomson (1989). "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment." Econometrica, 57, 447-480.

36. Mailath, G. and L. Samuelson (2001), "Who Wants a Good Reputation?" The Review of Economic Studies, 68, 415-441.

37. Mailath, G. and L. Samuelson (2006) Repeated Games and Reputations: Long-run Relationships. Oxford University Press, New York.

38. Medoff, J. and K. Abraham (1980). "Experience, Performance and Earnings", Quarterly Journal of Economics, 95, 703-736.

39. Mirman, L., L. Samuelson, and A. Urbano (1993). "Duopoly Signal Jamming," Economic Theory, Springer, 3(1), 129-49.

40. Obara, I. (2009). “Folk Theorem with Communication,” Journal of Economic Theory, 144, 120-134.

41. Phelan, C. (2006). "Public trust and government betrayal," Journal of Economic Theory, 130, 27-43.

42. Radner, R. (1985). “Repeated Principal-Agent Games with Discounting,” Econometrica, 53, 1173–1198.

43. Rahman, D. (2009). "But Who will Monitor the Monitor?," mimeo.

44. Renault, J. (2009). "Repeated Games with Incomplete Information", in Meyers, Robert (Ed.) Encyclopedia of Complexity and Systems Science, pp 7630-7651. Springer New York.

45. Rubinstein, A. (1979b). “An Optimal Conviction Policy for Offenses That May Have Been Committed byAccident,” in Applied Game Theory, ed. by S. J. Brams, A. Schotter, and G. Schwodiauer, 406–413. Physical-Verlag, Würzburg.

46. Rubinstein, A., and M. E. Yaari (1983). “Repeated Insurance Contracts and Moral Hazard,” Journal of Economic Theory, 30, 74–97.

47. Salanié, B. (2005) The Economics of Contracts: A Primer, 2nd Edition. MIT Press, Cambridge, MA.

48. Sannikov, Y. (2007). “Games with Imperfectly Observable Actions in Continuous Time,” Econometrica, 75, 1285–1329.

49. Shapiro, C. and J. Stiglitz (1984) "Equilibrium Unemployment as a Worker Discipline Device," American Economic Review, 74(3), 433-444.

50. Stiglitz, J. (1974). "Incentives and Risk Sharing in Sharecropping," The Review of Economic Studies 41 (2), 219-255.

51. Taylor, J. (1980). "Aggregate Dynamics and Staggered Contracts." Journal of Political Economy, 1, 1-23.

52. Tirole, J. (1986). "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations", Journal of Law, Economics and Organization, 2, 181-214.

53. Zheng, B. (2008). “Approximate Efficiency in Repeated Games with Correlated Private Signals,” Games and Economic Behavior, 63, 406-416.

3.9 *Appendix - Proofs of Lemmas, Propositions and Corollaries*

The proof of the various Lemmas makes extensive use of the dynamic programming methods developed by Abreu, Pearce and Stacchetti (1986, 1990).

Proof of Lemma 21. Consider player $i = a$. Suppose that in the beginning of the game, i.e. state 0, the agent selects an effort e^H, receiving an expected payoff π_a^H. In the end of the first period, if a low output is observed, the game enters in the punishment stage with expected payoff $\underline{v}_a / (1-\delta)$. This event occur with probability p^E. If a high output is observed, it might be uninformative in which case the value of the game for the agent associated with action e^H is $v_a^{H,L} / (1-\delta)$. Notice that we have a recursive pattern here, this case is equivalent to a repetition of the initial game one period later. This event occur with probability $1 - p^E - (1-\beta) r$.

With probability $(1-\beta) r$ the signal might be revealing, in which case the agent adjust his effort accordingly to e^L, obtaining an expected payoff of π_a^L for the following period. In this case there is no more learning, and we have a simple recursive structure. With probability q^T the agent is punished in the following period and with the remaining probability he obtains the value $\ddot{v}_a^{L} / (1-\delta)$.

Formally we have two recursive patterns, they are respectively

$$v_a^{H,L} = (1-\delta)\,\pi_a^H + \delta\left[p^E \underline{v}_a + (1-\beta)\, r \ddot{v}_a^{L} + \left(1 - p^E - (1-\beta)\, r\right) v_a^{H,L}\right],$$

and

$$\ddot{v}_a^{L} = (1-\delta)\,\pi_a^L + \delta\left[q^T \underline{v}_a + \left(1 - q^T\right) \ddot{v}_a^{L}\right],$$

which can be solved for $v_a^{H,L}$ to obtain (50). Reasoning in a similar way we obtain expression (50) for the principal. Employing the substitution suggested in the text we obtain the payoffs $v_i^{L,L}$ and $v_i^{H,H}$ for $i \in \{a, p\}$. ■

Proof of Proposition 23. First we search for the condition $\delta < \delta^+$. After observing an informative signal the agent chooses e^L if $\ddot{v}_a^{H} < \ddot{v}_a^{L}$. Where

$$\ddot{v}_a^{H} \equiv \frac{(1-\delta)\,\pi_a^H + \delta p^T \underline{v}_a}{1 - \delta\,(1-p^T)} \; \textit{and} \; \ddot{v}_a^{L} \equiv \frac{(1-\delta)\,\pi_a^L + \delta q^T \underline{v}_a}{1 - \delta\,(1-q^T)}, \tag{66}$$

with $v_a^{;;H}$ being the value of the infinitely repeated state 1 subgame for the informed agent when he chooses e^H and $v_a^{;;L}$ has a similar interpretation but with the agent choosing e^L. Rearranging for δ we obtain

$$\delta < \frac{\pi_a^L - \pi_a^H}{(\pi_a^L - \pi_a^H) + q^T(\pi_a^H - \underline{v}_a) - p^T(\pi_a^L - \underline{v}_a)} \equiv \delta^+. \tag{67}$$

Since $\pi_a^L > \pi_a^H$ and if $q^T\left(\pi_a^H - \underline{v}_a\right) > p^T\left(\pi_a^L - \underline{v}_a\right)$, i.e. condition (54) holds, we have $0 < \delta^+ < 1$.

The agent infinitely repeated game expected payoff when he supplies e^H while uninformed and e^L in the state 1 is $v_a^{H,L}$, and given by (50), the infinitely repeated game expected payoff when supplying e^L in any informational state is denoted and given by

$$v_a^{L,L} \equiv \frac{(1-\delta)\pi_a^L + \delta q^E \underline{v}_a}{D^L} + \delta(1-\beta)\frac{s}{D^L}\frac{(1-\delta)\pi_a^L + \delta q^T \underline{v}_a}{1-\delta(1-q^T)}, \tag{68}$$

where $D^L \equiv 1 - \delta\left(1 - q^E - (1-\beta)s\right)$. Solving $v_a^{H,L} \geq v_a^{L,L}$ for δ with equality we obtain the expression for δ^-. To guarantee the nonemptiness of the interval $\left[\delta^-, \delta^+\right)$, i.e. $\delta^+ > \delta^-$, we plug δ^+ in $v_a^{H,L}$ and $v_a^{L,L}$. After some algebra simplifications the inequality relation $v_a^{H,L} > v_a^{L,L}$ becomes

$$\frac{r\beta\left(\pi_a^H - \pi_a^L\right)\left(q^T\left(\pi_a^H - \underline{v}_a\right) - p^T\left(\pi_a^L - \underline{v}_a\right)\right)}{r(\pi_a^L - \pi_a^H) - p^T(\pi_a^H - \underline{v}_a) + q^T(\pi_a^H - \underline{v}_a)} > \frac{\beta s\left(\pi_a^H - \pi_a^L\right)\left(q^T\left(\pi_a^H - \underline{v}_a\right) - p^T\left(\pi_a^L - \underline{v}_a\right)\right)}{s(\pi_a^L - \pi_a^H) - p^T(\pi_a^L - \underline{v}_a) + q^T(\pi_a^L - \underline{v}_a)},$$

further manipulations lead us to

$$\left(\frac{s}{(\pi_a^L - \pi_a^H)s + (\pi_a^L - \underline{v}_a)(q^T - p^T)} - \frac{r}{(\pi_a^L - \pi_a^H)r + (\pi_a^H - \underline{v}_a)(q^T - p^T)}\right)$$
$$\times\beta\left(\pi_a^L - \pi_a^H\right)\left(q^T\left(\pi_a^H - \underline{v}_a\right) - p^T\left(\pi_a^L - \underline{v}_a\right)\right) > 0.$$

From which we require conditions (54) and (55) to hold simultaneously. Since $\pi_a^L > \pi_a^H > \underline{v}_a$, it is easy to show that $\delta^- > 0$. Putting all together conditions (54) and (55) guarantee that $0 < \delta^- < \delta^+ < 1$. ■

Proof of Corollary 25. The reference payoff $v_i^{H,H}$ is given by (56). The expression for $v_i^{H,L}$ is given by (50) in Lemma 21.

(i) The conditions of Proposition 23 establish that $v_a^{H,L} > v_a^{H,H}$ for $\delta < \delta^+$ where δ^+ solves $v_a^{H,L} = v_a^{H,H}$.

(ii) Solving $v_p^{H,L} \leq v_p^{H,H}$ for δ and using the fact that $\pi_p^L = \underline{v}_p$ we obtain

$$\delta \leq \frac{1}{1+q^T} \equiv \delta^{ref}. \tag{69}$$

It enough to show that $\delta^+ \leq \delta^{ref}$; it lead us to the following condition

$$0 \leq \left(\pi_a^L - \underline{v}_a\right)\left(q^T - p^T\right),$$

which is satisfied with strict inequality, since the right-hand side is strictly positive, i.e. $\pi_a^L > \underline{v}_a$ and $q^T > p^T$. Then $v_p^{H,L} < v_p^{H,H}$ for all $\delta < \delta^+$.

(iii) Solving $v_a^{H,H} + v_p^{H,H} > v_a^{H,L} + v_p^{H,L}$ for δ we obtain three roots $\delta > 0$, $\delta < 1$ and

$$\delta > \frac{\left(\pi_p^H - \underline{v}_p\right) - \left(\pi_a^L - \pi_a^H\right)}{\left(\pi_a^H + \pi_p^H - \underline{v}_p\right)\left(1-q^T\right) - \pi_a^L\left(1-p^T\right) + \underline{v}_a\left(q^T - p^T\right)}.$$

The third root is larger than one if

$$q^T\left(\pi_a^H - \underline{v}_a\right) + q^T\left(\pi_p^H - \underline{v}_p\right) > p^T\left(\pi_a^L - \underline{v}_a\right),$$

which is guaranteed by condition (54). ■

Proof of Proposition 26. The problem is equivalent to a infinitely repeated game without learning and the monitor type known to be θ^E. In this case the players payoffs are simply

$$v_i^{H*} = \frac{(1-\delta)\,\pi_i^H + \delta p^E \underline{v}_i}{1-\delta\left(1-p^E\right)} \text{ and } v_i^{L*} = \frac{(1-\delta)\,\pi_i^L + \delta q^E \underline{v}_i}{1-\delta\left(1-q^E\right)}, \tag{70}$$

when the agent provides high and low effort respectively and $i \in \{a, p\}$. The recursive structure is simple; the agent supplies a given effort and with probability p^E or q^E he is punished or with the remaining probabilities the same pattern is repeated. Solving $v_a^{H*} \geq v_a^{L*}$ for δ we obtain

$$\delta \geq \frac{\pi_a^L - \pi_a^H}{\left(\pi_a^L - \pi_a^H\right) + \left(\pi_a^H - \underline{v}_a\right)q^E - \left(\pi_a^L - \underline{v}_a\right)p^E} \equiv \delta^*.$$

Then $\delta^* \in (0,1)$ if $q^E\left(\pi_a^H - \underline{v}_a\right) > p^E\left(\pi_a^L - \underline{v}_a\right)$, which is a general version of condition (54) for all $\beta \in (0,1)$. Since p^E and q^E depends linearly on β, and the sum of (54) and (55) gives $q^S\left(\pi_a^H - \underline{v}_a\right) > p^S\left(\pi_a^L - \underline{v}_a\right)$, the lowest difference $q^E\left(\pi_a^H - \underline{v}_a\right) - p^E\left(\pi_a^L - \underline{v}_a\right)$ must be reached at $\beta \to 0$ by condition (55). This difference is always positive. The value of δ^+ is given by (67) and is larger than δ^* when (55) holds.

To show that $\delta^* < \delta^-$, just replace δ^* for δ in $v_a^{H,L} \leq v_a^{L,L}$ and rearrange to obtain

$$(1-\beta)\left(\pi_a^L - \pi_a^H\right)^2 \left(s\left(\pi_a^H - \underline{v}_a\right) - r\left(\pi_a^L - \underline{v}_a\right)\right)\left(q^E\left(\pi_a^H - \underline{v}_a\right) - p^E\left(\pi_a^L - \underline{v}_a\right)\right) > 0,$$

which, following the previous argument, is satisfied when both (54) and (55) hold. ■

Proof of Proposition 27. Manipulate the inequality $v_i^{H,H} > v_i^{H*}$, which the expressions are given by (56) and (70) respectively, we obtain

$$r(1-\beta)\left(p^E - p^T\right)\left(\pi_i^H - \underline{v}_i\right)(1-\delta)\delta^2 > 0,$$

which is clearly larger than zero for all $\delta \in (0,1)$, since $p^E = p^T + \beta r > p^T$, $\pi_i^H > \underline{v}_i$ for $i \in \{a, p\}$ and $\beta \in (0,1)$. Participation is guaranteed if $v_i^{H*} \geq \underline{v}_i$. ■

Proof of Corollary 28. Lets start noticing that replacement cost are only incurred by the principal. (i) For $\delta \in \left[\delta^+, 1\right)$ the market compensation is sufficient to keep the agent with incentives. The principal never replaces the monitor.

(ii) By Proposition 26, without replacement costs and with full freedom in the monitor replacement the principal cannot obtain more than v_p^{H*} then she must not be able to do more if we introduce replacement costs. In this case the agent obtain v_a^{H*} as his worst payoff. On the other hand if costs are too high and in consequence payoffs are expected to fall bellow $v_p^{H,L}$, then principal has always the option to never replace the monitor, as in Proposition 23, guaranteeing at least a payoff of $v_p^{H,L}$. In this case the agent benefit from the principal replacement passivity and obtain the payoff $v_a^{H,L}$.

(iii) Without replacement strategies both player would obtain $\underline{v}_p$ and $\underline{v}_a$. Costless replacement strategies expand the discount region to δ^*. If replacement strategies improve over $v_p^{L,L} = \underline{v}_p$ then the principal must obtain a payoff of at most v_p^{H*}. When replacement cost needed to provide the agent with incentives are too expensive, i.e. $v_p \leq \underline{v}_p$, the principal sticks to her outside option. The agent cannot get less than v_a^{H*} when the principal employs replacement strategies. Only if the principal finds optimal to not participate, in this case he obtains $\underline{v}_a$. Providing that the replacement costs are not too high the principal will employ replacement strategies. In order to keep the principal with incentives the agent has to provide high effort at least when uninformed, in the best scenario he would obtain $v_a^{H,L}$. ■

Proof of Lemma 29. Consider the case where the principal change the monitor as soon as an informative signal is observed, that is $v_{i,k,0}^{H,L}$. Using a similarly reasoning used to prove Lemma 21, we obtain the following recursive payoffs for the case where $i = p$,

$$v_{p,k,0}^{H,L} = (1-\delta)\left(\pi_p^H - k\right) + \delta\left[p^E \underline{v}_p + (1-\beta)\, r v_{p,k,0}^{H,L} + \left(1 - p^E - (1-\beta)\, r\right) v_{p,0,0}^{H,L}\right],$$

and

$$v_{p,0,0}^{H,L} \equiv (1-\delta)\,\pi_p^H + \delta\left[p^E \underline{v}_p + (1-\beta)\, r v_{p,k,0}^{H,L} + \left(1 - p^E - (1-\beta)\, r\right) v_{p,0,0}^{H,L}\right].$$

The first expression $v_{p,k,0}^{H,L}$ is the value of the repeated game that starts with a costly replacement of the monitor. The agent starts providing high effort, then in the following period he is punished with probability p^E. A revealing signal occurs with probability $(1-\beta)\, r$, in which case we the principal replace immediately the monitor and we have a recursive structure. With probability the $1 - p^E - (1-\beta)\, r$ none of these events occur and the monitor is not replaced. The second expression $v_{p,0,0}^{H,L}$ is the value of the relation in this case, which has a recursive structure. Solving recursively these two expression for $v_{p,k,0}^{H,L}$ we obtain expression (57) for the case where $n = 0$. Notice that we are solving the recursion assuming that there is a replacement cost in the beginning of the game. This simplifies the recursion, to correct it in the end we add $k\,(1-\delta)$.

Consider now the case where the principal replace the monitor one period after a revealing signal is observed, in this case we have the following system of equations

$$\begin{aligned}
v_{p,k,1}^{H,L} &= (1-\delta)\left(\pi_p^H - k\right) + \delta\left[p^E \underline{v}_p + (1-\beta)\, r v_{p,0,1}^{\cdot,L} + \left(1 - p^E - (1-\beta)\, r\right) v_{p,0,1}^{H,L}\right], \\
v_{p,0,1}^{H,L} &\equiv (1-\delta)\,\pi_p^H + \delta\left[p^E \underline{v}_p + (1-\beta)\, r v_{p,0,1}^{\cdot,L} + \left(1 - p^E - (1-\beta)\, r\right) v_{p,0,1}^{H,L}\right], \\
v_{p,0,1}^{\cdot,L} &\equiv (1-\delta)\,\pi_p^L + \delta\left[q^T \underline{v}_p + \left(1 - q^T\right) v_{p,k,1}^{H,L}\right].
\end{aligned}$$

Notice that now we have a third equation, that is due to the fact that the agent is allowed to enjoy the benefit of learning for one period. After we have solve for $v_{p,k,1}^{H,L}$ and added $k\,(1-\delta)$ we obtain expression (57) for the case where $n = 1$. For general n, in the previous system of equations replace the value $v_{p,k,1}^{H,L}$ for $v_{p,k,n}^{H,L}$, $v_{p,0,1}^{H,L}$ for $v_{p,0,n}^{H,L}$ and $v_{p,0,1}^{\cdot,L}$ for $v_{p,0,n}^{\cdot,L}$ in the two first equations and substitute the third equation by

$$v_{p,0,n}^{\cdot,L} = \left((1-\delta)\,\pi_p^L + \delta q^T \underline{v}_p\right) \sum_{l=0}^{n-1} \delta^l (1-q^T)^l + \delta^n \left(1-q^T\right)^n v_{p,k,n}^{H,L}. \tag{71}$$

Solve the system for $v_{p,k,n}^{H,L}$, using the fact that $\sum_{l=0}^{n-1} x^l = (1-x^n)/(1-x)$ and adding the term $k(1-\delta)$, we obtain expression (57). A similar reasoning is done when $i = a$ with $k = 0$. ■

Proof of Proposition 30. First we show that $n = 0$ is an optimal choice when $v_{p,k,n}^{H,L} > v_{p,k,n+1}^{H,L}$, i.e. $k < k^{cut}$, and $n \to \infty$ is optimal if $v_{p,k,n}^{H,L} \leq v_{p,k,n+1}^{H,L}$, i.e. $k \geq k^{cut}$. The value k^{cut} is the value of k that solves $v_{p,k,n}^{H,L} = v_{p,k,n+1}^{H,L}$. Using (57) for n and $n+1$, setting $\pi_p^L = \underline{v}_p$, after some algebraic manipulation we obtain k^{cut} given in (62). Notice that k^{cut} is independent of n. Clearly $k^{cut} > 0$ since both the numerator and denominator are strictly positive. Then $v_{p,k,0}^{H,L}$ is the supremum of $\left\{v_{p,k,0}^{H,L}, v_{p,k,1}^{H,L}, \ldots\right\}$ when $k < k^{cut}$ and $v_{p,k,\infty}^{H,L}$ is the supremum when $k \geq k^{cut}$.

Lets first look at the case where $\delta \in \left[\delta^-, \delta^+\right)$. In order for $n = 0$ to be optimal the principal must obtain a payoff larger than $v_p^{H,L} = v_{p,k,\infty}^{H,L}$. This is not a participation constraint, but a condition for the monitor replacement. Solving $v_{p,k,n}^{H,L} > v_p^{H,L}$ for k, we again obtain (62). Then $n = 0$ is optimal if $k \in \left[0, k^{cut}\right)$, otherwise $n \to \infty$ is optimal, i.e. for $k \geq k^{cut}$. The agent incentives are guaranteed by Proposition 23 and since $\left[\delta^-, \delta^+\right) = \left[\delta_\infty^-, \delta_\infty^+\right) \subset \left[\delta_0^-, \delta_0^+\right)$ by (60) and (61).

In order to find the principal optimal strategy for $\delta \in \left[\delta^*, \delta^-\right)$ both the principal and the agent must have incentives in participate. Recall that $\pi_p^L = \underline{v}_p$, imposes a lower bound on low effort, see (53). The principal participation for this discounting region is guaranteed if $v_{p,k,n}^{H,L} > v_p^{L,L} = \underline{v}_p$. Solving the inequality for k we obtain $k < k^{PCn}$, where k^{PCn} is given by (63). Moreover, k^{PCn} is a strictly positive and increasing function of n since $\delta\left(1-q^T\right) < 1$ and $\pi_p^H > \underline{v}_p$, i.e. $0 < k^{PCn} < k^{PCn+1}$ for all $n \in \mathbb{N}_0$. To show that $k^{PCn} > k^{cut}$, notice that k^{PCn} reach its lowest value when $n = 0$. It is then the hardest to satisfy scenario, but in this case $1 - \delta\left(1 - p^T - \beta r\right) > 0$, implying that $k^{cut} < k^{PCn}$ for all $n \in \mathbb{N}_0$.

Recall that when $k < k^{cut}$, the principal participation is guaranteed for all n and $n = 0$ is the optimal replacement choice, this is also true for $\delta \in \left[\delta^*, \delta^-\right)$. The ordering of sequence

(61) guarantees that the agent incentives for e^H in any uninformed state are satisfied. Consider now that $k \geq k^{cut}$, in this case the principal's optimal choice has to guarantee that the agent incentives to provide at least high effort in the uninformed state are satisfied. When $\delta \in \left[\delta_n^-, \delta_{n+1}^-\right) \subset \left[\delta^*, \delta^-\right)$ the principal optimal choice must be n. A choice of $n-1$ is payoff inferior for the principal since $k \geq k^{cut}$, a choice $n+1$ does not provide the agent with incentives. Finally the principal participation has also to be guaranteed, i.e. $k < k^{PCn}$, otherwise $v_{p,k,n}^{H,L} \leq \underline{v}_p$ and no trade is optimal. ■

Proof of Corollary 31. Since in Proposition 30 the principal participation constraints are on same time replacement constraints any finite optimal choice that allows for trade must allow a payoff improvement, the proof of this result can be found on the proof of Proposition 30. See also the proof of Corollary 28. ■

Proof of Lemma 32. The proof is identical to the proofs of Lemmas 21 and 29. We look first for $\widetilde{v}_{p,k,0}^{H,L}$, i.e. the case where the agent is substituted in every period. Let $i = p$, we obtain the following recursive payoff

$$\widetilde{v}_{p,k,0}^{H,L} = (1-\delta)\left(\pi_p^H - k\right) + \delta\left[p^E \underline{v}_p + (1-\beta)\, r \widetilde{v}_{p,k,0}^{H,L} + \left(1 - p^E - (1-\beta)\, r\right)\widetilde{v}_{p,k,0}^{H,L}\right],$$

which can be solved for $\widetilde{v}_{p,k,0}^{H,L}$ and finally add $k(1-\delta)$. Consider now the case where $n = 1$, in this case we obtain the following system of equations

$$\widetilde{v}_{p,k,1}^{H,L} = (1-\delta)\left(\pi_p^H - k\right) + \delta\left[p^E \underline{v}_p + (1-\beta)\, r \widetilde{v}_{p,0,1}^{\cdot,L} + \left(1 - p^E - (1-\beta)\, r\right)\widetilde{v}_{p,0,1}^{H,\cdot}\right],$$

$$\widetilde{v}_{p,0,1}^{\cdot,L} = (1-\delta)\,\pi_p^L + \delta q^T \underline{v}_p + \delta\left(1 - q^T\right)\widetilde{v}_{p,k,1}^{H,L},$$

$$\widetilde{v}_{p,0,1}^{H,\cdot} = (1-\delta)\,\pi_p^H + \delta p^E \underline{v}_p + \delta\left(1 - p^E\right)\widetilde{v}_{p,k,1}^{H,L},$$

which is solved for $\widetilde{v}_{p,k,1}^{H,L}$ and finally adding $k(1-\delta)$. In this case the agent is substituted after two periods. The first equation is the value of the repeated game that starts with a costly selection of a monitor. The agent starts providing high effort. In the end of the first period he is punished with probability p^E, otherwise either a revealing signal occurs with probability $(1-\beta)\, r$ or he stays uninformed with probability $1 - p^E - (1-\beta)\, r$. In the former case he enjoys the benefits from learning during on period after which he is replaced in case that is not punished. The repeated game restarts again with a costly replacement of the monitor. This is the second equation which is the value of the relation in the case where the agent gets informed and supplies with low effort for one period. The third expression has a similar interpretation as the second one, but for the case where the game remains in the uninformed state, the agent start by supplying high effort. If no punishment occurs, with probability $1-p^E$ the repeated game is restarted again.

Consider now the general case n, in this case we have $2n+1$ equations,

$$\begin{aligned}
\widetilde{v}_{p,k,n}^{H,L} &= (1-\delta)\left(\pi_p^H - k\right) + \delta\left[p^E \underline{v}_p + (1-\beta)\, r \widetilde{v}_{p,0,n}^{\cdot,L} + \left(1-p^E-(1-\beta)\, r\right)\widetilde{v}_{p,0,n}^{H,\cdot}\right], \\
\widetilde{v}_{p,0,n}^{\cdot,L} &= \left((1-\delta)\,\pi_p^L + \delta q^T \underline{v}_p\right)\sum_{k=0}^{n-1}\delta^k\left(1-q^T\right)^k + \delta^n\left(1-q^T\right)^n \widetilde{v}_{p,k,n}^{H,L}, \\
\widetilde{v}_{p,0,n}^{H,\cdot} &= (1-\delta)\,\pi_p^H + \delta\left[p^E \underline{v}_p + (1-\beta)\, r \widetilde{v}_{p,0,n-1}^{\cdot,L} + \left(1-p^E-(1-\beta)\, r\right)\widetilde{v}_{p,0,n-1}^{H,\cdot}\right], \\
&\quad\ldots \\
\widetilde{v}_{p,0,2}^{\cdot,L} &= \widetilde{v}_p^{0,2\prime} = \left((1-\delta)\,\pi_p^L + \delta q^T \underline{v}_p\right)\left(1+\delta\left(1-q^T\right)\right) + \delta^2\left(1-q^T\right)^2\widetilde{v}_{p,k,n}^{H,L}, \\
\widetilde{v}_{p,0,2}^{H,\cdot} &= \widetilde{v}_p^{0,2} = (1-\delta)\,\pi_p^H + \delta\left[p^E \underline{v}_p + (1-\beta)\, r \widetilde{v}_{p,0,1}^{\cdot,L} + \left(1-p^E-(1-\beta)\, r\right)\widetilde{v}_{p,0,1}^{H,\cdot}\right], \\
\widetilde{v}_{p,0,1}^{\cdot,L} &= (1-\delta)\,\pi_p^L + \delta q^T \underline{v}_p + \delta\left(1-q^T\right)\widetilde{v}_{p,k,n}^{H,L}, \\
\widetilde{v}_{p,0,1}^{H,\cdot} &== (1-\delta)\,\pi_p^H + \delta p^E \underline{v}_p + \delta\left(1-p^E\right)\widetilde{v}_{p,k,n}^{H,L},
\end{aligned}$$

which can be solved recursively for $\widetilde{v}_{p,k,n}^{H,L}$ to obtain, after adding $k\,(1-\delta)$, the expression

$$\widetilde{v}_{p,k,n}^{H,L} = \frac{\left((1-\delta)\pi_p^H + \delta p^E \underline{v}_p\right)\sum_{k=0}^{n}\delta^k z^k + \left((1-\delta)\pi_p^L + \delta q^T \underline{v}_p\right)\sum_{r=1}^{n} y z^{r-1}\sum_{k=0}^{n-r}\delta^{k+r}x^k - k(1-\delta)}{1-\delta^{n+1}\left(z^{n+1}+\sum_{k=0}^{n} x^{n-k} y z^k\right)} + k\,(1-\delta)\,, \quad (72)$$

which equals expression (64) after all the summations have been solved. Similar reasoning is employed when $i=a$ with $k=0$. ■

When required to shorten in notation we use the following definitions; $x \equiv 1-q^T$, $y \equiv (1-\beta)\, r$ and $z \equiv 1-p^E-(1-\beta)\, r = 1-p^T-r$.

Proof of Proposition 33. Suppose that $\widetilde{v}_{p,k,n}^{H,L}$ has a unique global maximum in $\mathbb{N}_0$. Then $\widetilde{v}_{p,k,n}^{H,L}$ given by (64) reach its maximum value at n if $\widetilde{v}_{p,k,n-1}^{H,L} \leq \widetilde{v}_{p,k,n}^{H,L}$ and $\widetilde{v}_{p,k,n}^{H,L} \geq \widetilde{v}_{p,k,n+1}^{H,L}$. Let $k^{n,n+1}$ be the value k that solves $\widetilde{v}_{p,k,n}^{H,L} = \widetilde{v}_{p,k,n+1}^{H,L}$ for $n=0,1,2....$ Consider $\pi_p^L = \underline{v}_p$, after some algebraic manipulation we obtain

$$k^{n,n+1} = \frac{\left(\pi_p^H - \underline{v}_p\right) y\left[z^{n+1}(1-\delta z) - x^{n+1}(1-\delta x) + z^{n+1}x^{n+1}\delta^{n+1}(z-x)\,\delta\right]}{(1-\delta z)\left[z^{n+1}(1-\delta z)(z+y-x) - x^{n+1}y\,(1-\delta x)\right]},$$

Notice that $k^{-1,0}=0$. The expression for $k^{n,n+1}$ is increasing in n if

$$y \leq \frac{(x-z)\left(1-\delta^{n+1}z^{n+1}\right)}{\delta^{n+1}\left(x^{n+1}-z^{n+1}\right)}. \tag{73}$$

Independently of the relation between x and z, the right-hand side is monotonically increasing in n, converging to ∞ when $n \to \infty$ and to $(1-\delta z)/\delta$ when $n \to 0$. The latter it is the lowest value and for that reason the hardest to satisfy. Then substituting $y \equiv (1-\beta)\,r$ and $z \equiv 1-p^T-r$ and solving for δ we obtain $\delta \leq 1/\left(1-p^T-\beta r\right)$, which is always satisfied because its larger than 1. Then the optimal replacement choice n has to satisfy $\widetilde{v}^{H,L}_{p,k,n-1} \leq \widetilde{v}^{H,L}_{p,k,n}$ and $\widetilde{v}^{H,L}_{p,k,n} \geq \widetilde{v}^{H,L}_{p,k,n+1}$, or equivalently $k^{n-1,n} \leq k \leq k^{n,n+1}$.

Now, let $n \to \infty$ in the expression for $k^{n,n+1}$. If $x \geq z$, i.e. $r \geq q^T - p^T$, then we obtain

$$k^{\infty} = \frac{\pi^H_p - \underline{v}_p}{1-\delta z} = \frac{\pi^H_p - \underline{v}_p}{1-\delta\left(1-p^T-r\right)},$$

which is expression (62). When $x < z$, i.e. $r < q^T - p^T$, we obtain

$$k^{\infty} = \frac{\pi^H_p - \underline{v}_p}{1-\delta z}\frac{y}{z+y-x} = \frac{\pi^H_p - \underline{v}_p}{1-\delta\left(1-p^T-r\right)}\frac{(1-\beta)\,r}{q^T-p^T-\beta r},$$

where the second ratio on the right-hand side is smaller than one. Also, since $p^T + r = p^S < 1$, we have $k^{\infty} > 0$ always. In resume for any $\delta \in \left[\delta^*, \delta^+\right)$ and $k \in [0, k^{\infty})$, when unrestricted the principal prefers to replace the monitor after a finite number of periods, while if $k \geq k^{\infty}$ he must choose $n \to \infty$.

We need now to verify when the principal prefers to employ replacement strategies. For this $\widetilde{v}^{H,L}_{p,k,n} \geq \underline{v}_p$ when $\delta \in \left[\delta^*, \delta^-\right)$ and $\widetilde{v}^{H,L}_{p,k,n} \geq v^{H,L}_p$ when $\delta \in \left[\delta^-, \delta^+\right)$. Lets look first to the former case. Solve $\widetilde{v}^{H,L}_{p,k,n} = \underline{v}_p$ for k to obtain the participation condition in terms of replacement costs, denote it by k^{PCn} and is given by

$$k^{PCn} = \frac{\left(\pi^H_p - \underline{v}_p\right)(x-z)\left(1-\delta^{n+1}z^{n+1}\right)}{\left[\left(x^{n+1}-z^{n+1}\right)y + z^{n+1}(x-z)\right](1-\delta z)}.$$

Independently of the relation between x and z, the function k^{PCn} is strictly increases in n. Formally it is enough to show that $k^{PCn} < k^{PCn+1}$, that is

$$y > -\frac{(x-z)\,z^{n+1}\,(1-\delta z)}{\left(x^{n+1}-z^{n+1}\right)-\delta\left(x^{n+2}-z^{n+2}\right)+\delta\,(x-z)\,\delta^{n+1}x^{n+1}z^{n+1}}. \tag{74}$$

The right-hand side is always negative, since $\left(x^{n+1}-z^{n+1}\right)>\left(x^{n+2}-z^{n+2}\right)$ when $x>z$ and the reverse when $z>x$. The numerator and denominator always have the same signal. Then $k^{PC\infty}$ is the largest cost and takes the value ∞ when $n\to\infty$.

Consider now the case $\delta\in\left[\delta^{*},\delta^{-}\right)$, and solve $\widetilde{v}_{p,k,n}^{H,L}=v_{p}^{H,L}$ for k to obtain the replacement condition

$$k^{RCn}=\frac{\left(\pi_{p}^{H}-\underline{v}_{p}\right)y\left(x^{n+1}-z^{n+1}\right)}{\left[\left(x^{n+1}-z^{n+1}\right)y+\left(x-z\right)z^{n+1}\right]\left(1-\delta z\right)}.$$

Taking the limit $n\to\infty$ we found that $k^{RCn}\to k^{\infty}$ whether $x\geq z$ or $x<z$. Moreover k^{RCn} is strictly increasing with n. In order for $k^{RCn}>k^{n,n+1}$, the same condition (73) must be strictly satisfied, which we have shown above to be always the case. Moreover, since $v_{p}^{H,L}>\underline{v}_{p}$ then $k^{PCn}>k^{RCn}$, consequently we have $k^{PCn}>k^{n,n+1}$ for all $n\in\mathbb{N}_{0}$.

Now we consider the agent incentives to place high effort in the uniformed state, given a replacement choice. (i) For $\delta\in\left[\delta^{-},\delta^{+}\right)$, since for all n, $\delta_{n}^{-}\leq\delta^{-}$ the agent incentives to provide high effort while uninformed are always satisfied. The principal choose some finite n if $k\in[0,k^{\infty})$ and $n\to\infty$ otherwise. However if $k\geq k^{RCn}$, she must choose the smallest τ larger than n such that the inequality is reversed, i.e. $k<k^{RC\tau}$. The monitor has to be replaced less often in order to increase the replacement condition bound $k^{RC\tau}$. However, if such it is not possible, i.e. $k\geq k^{RC\infty}=k^{\infty}$, the principal should choose $n\to\infty$. Since $k^{PC\infty}=\infty$, participation is always guaranteed.

Now suppose that $\delta\in\left[\delta_{m}^{-},\delta_{m+1}^{-}\right)\subseteq\left[\delta^{*},\delta^{-}\right)$. Then we have two possibilities. (ii) The unconstrained optimal $n\leq m$, in this case the choice n is optimal if $k<\ k^{PCn}<k^{\infty}$, since the principal and agent incentives are met. However, if $k^{PCn}\leq k<k^{\infty}$, the principal has to choose the lowest $\tau\in\{n+1,...,m\}$ such that $k<k^{PC\tau}$. The principal cannot choose a replacement frequency above m because in this case the agent would not have incentives to provide effort in the uninformed state. Similarly, when $k>k^{\infty}$, she must choose m if $k<k^{PCm}$. It might happen that $k\geq k^{PCm}$ then the principal should not participate because the replacement costs are too high and for any replacement frequency above m the agent has no incentives. (iii) Second, if the unconstrained optimal $n>m$, to keep the agent with incentives, it is optimal to choose m, but such is only the case if $k<k^{PCm}$. Otherwise, trade is not possible, because it is too costly for the principal to provide the agent with

incentives. ■

Proof of Proposition 34. The argument is similar to the one used in the Proof of Corollary 31. In Proposition 33 the principal participation constraints are on same time replacement constraints. Then any finite optimal choice that allows for trade must allow a payoff improvement, the proof of this result can be found along the proof of Proposition 33. ■